

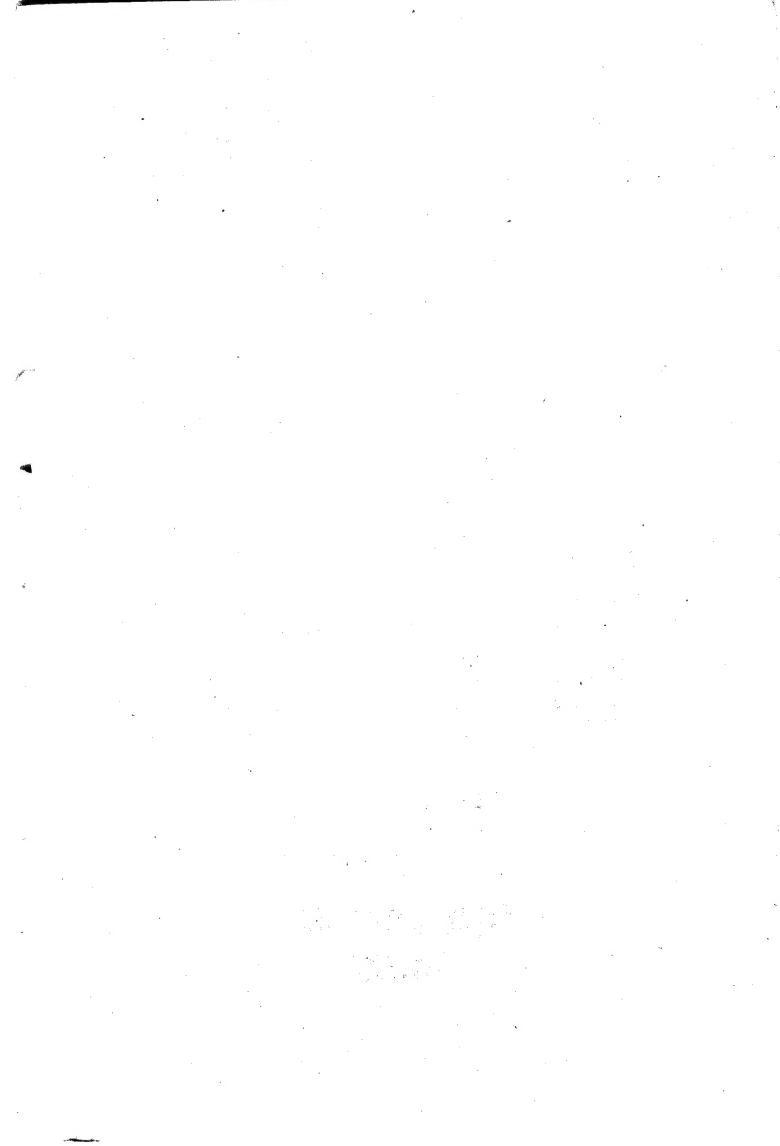
الإحصاء التطبيقي وبحوث العمليات

د/ عبد اللطيف عبد الفتاح أبو العلا / د/ إبراهيم محمد مهدي
أستاذ الإحصاء التطبيقي / أستاذ الإحصاء الإكتواري

د/ سلطان محمد عبد الحميد / د/ ياسر محمد العدل
أستاذ ورئيس قسم الإحصاء والتأمين / مدرس الإحصاء

٢٠٠٣ - ٢٠٠٤

مكتبة الجلاء الجديدة
المنصورة



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

وقل رب زدني علما

صدق الله العظيم

الجزء الأول

الإحصاء التطبيقي

أ.د/ عبد اللطيف عبد الفتاح أبو العلا

أ.د/ سلطان محمد عبد الحميد

مقدمة

تأتى هذه الطبعة فى إطار ما سبق أن اصدر من طبعات فى مجال الإحصاء التطبيقى والاستنتاج الإحصائى. ولكن ما يميز هذه الطبعة عن سابقتها هو الاتجاه إلى تقديم بعض الأساليب الإحصائية التطبيقية فى مجال التنبؤ واتخاذ القرارات فى ظل عدم اليقين.

وبالرغم من محدودية حجم هذه الطبعة إلا أنه روعى فى اختيار الموضوعات التى تناولتها التدرج والتكامل وإلا يكون الإيجاز فى العرض على حساب الأساسيات ، من أجل ذلك اختير محتوى هذه الطبعة ليشمل الفصول التالية:

الفصل الأول: تحليل التباين وتصميم التجارب .

الفصل الثانى: التصنيف متعدد الاتجاهات.

الفصل الثالث: تحليل الانحدار البسيط.

الفصل الرابع: تحليل الانحدار المتعدد.

الفصل الخامس: الطرق اللامعلمية.

وكان الهدف من الفصول الأربعة الأولى هو تحقيق انتشاع الزوية عند اتخاذ القرار سواء على أساس تعدد العينات أو تعدد المتغيرات أو بتوفير ظروف متشابهة للمتغيرات عند بدء التجربة أو تحريرها من أثر عدم التجانس السابق عند جمع المعلومات. أما الفصل الخامس والأخير ، فقد قدمنا بعض الأساليب اللامعلمية (أو اللابارامترية) فى اتخاذ القرارات ، وهى مجالات وأساليب حديثة ومتنوعة ولكن اختيار بعضها وتقديمه فى هذا المرجع - أن أمكن تجاوزاً أن يطلق عليه ذلك - جاء لتحقيق التناظر والتقابل مع ما سبق تقديمه من أساليب بارامترية كأدوات لاتخاذ القرارات أو حين تحول طبيعة القياسات دون استخدامها.

وأنه وإن كان الهدف أساساً من إعداد هذا الكتاب هو الدارس والباحث
فى المجالات التجارية والاقتصادية وهو ما قد تعكسه العديد من الأمثلة
والتطبيقات التى حفلت بها هذه الطبعة ، إلا أن الحاجة إلى هذه الأساليب
الإحصائية للدارس والباحث فى مجالات العلوم التطبيقية والإنسانية كالعلوم
الزراعية والتربية والاجتماع بل وبعض الدراسات الطبية والصحية لا تقل الحاحا
عنها فى مجالات الدراسات والبحوث الإدارية والاقتصادية فليس أيسر من إحلال
متغير محل آخر بما يتفق مع مجالات التطبيق. وهذا ما يميز الإحصاء كأسلوب
رقمى للقياس والتحليل والتشخيص عن العلوم الأخرى.
ولعلنا نكون قد وفقنا بفضل من الله فى تحقيق بعض ما كنا نهدف إليه
من إعداد لهذا المرجع المحدود.

والسلام عليكم ورحمة الله وبركاته،

أ.د/ عبد الطيف عبد الفتاح أبو العلا
أ.د/ سلطان محمد عبد الحميد

المنصورة فى أكتوبر ٢٠٠٣

الفصل الأول

تحليل التباين وتصميم التجارب Analysis of Variance & Design of Experiments

محتويات الفصل:

أولاً : تحليل التباين:

(١) اختبارات الفروض بشأن تباين مجتمع ما أو عدة مجتمعات ،
توزيع كاي^٢.

(١- ١) اختبارات الفروض بشأن تباين المجتمع σ^2 .

(٢-١) اختبارات الفروض بشأن تباين مجتمعين : $\sigma^2_1 = \sigma^2_2$ ،
توزيع ف (F).

(٢) اختبارات الفروض بشأن $(\mu_1 - \mu_2)$ أو عدة متوسطات .

(٣) تحليل التباين.

(٤) ملاحظات ختامية.

ثانياً : تصميم التجارب:

(١) تعاريف.

(١- ١) التجربة.

(٢- ١) المعالجات.

(٣- ١) وحدة التجربة.

(٤- ١) خطأ التجربة.

(٥- ١) العشوائية.

ثالثاً : التصميم كامل المشوائية

(١) استخدام عدد متساو من وحدات التجربة في كل معالجة.

(١- ١) النموذج الرياضي.

(١-٢) النموذج الحسابي والتحليل.

(٢) التصميم كامل العشوائية : عدد غير متساو من وحدات التجربة لكل معالجة.

(٣) العلاقة بين التصميم كامل العشوائية حيث ($l = 2$) واختبار

الفرض : $\mu_1 = \mu_2$

(٤) المقارنات الفردية.

(٥) تعليق ختامي.

تمارين.

الفصل الأول

تحليل التباين وتصميم التجارب

مقدمة: سوف نبدأ بمعالجة كيفية اختبار تباين مجتمع ما σ^2 وكذلك تباينات عدة مجتمعات مستقلة σ^2 وتقديم أدوات الاختبار التي تستخدم وهي توزيعات كاي² وكذلك نسبة التباين ف كنتمهيد لموضوع تحليل التباين.

أولاً : تحليل التباين

Analysis of Variance

(1) اختبارات الفروض بشأن تباين مجتمع ما أو عدة مجتمعات مستقلة.

تعرضنا في مرحلة سابقة إلى اختبارات الفروض الإحصائية المتعلقة بالوسط الحسابي لمجتمع ما μ باستخدام أدوات اختبار إحصائية تربط بين μ وتقديرها \bar{x} من عينة احتمالية أو عشوائية وذلك باستخدام Y (التوزيع المعتاد المعياري) أو T (توزيع t) بشرط توافر شروط معينة . وكذلك لاختبار الفرق بين متوسطين ($\mu_1 - \mu_2$) اعتماداً على تقديراتهما \bar{x}_1, \bar{x}_2 باستخدام اختبار Y للتوزيع المعتاد أو اعتماداً على توزيع T متى توافرت شروط معينة . كما تعرضنا أيضاً لاختبار نسبة متغير ما في المجتمع L أو الفرق بين نسبتي L_1, L_2 باستخدام تقديراتهما من عينات عشوائية من المجتمع أو المجتمعين أي $L_1 - L_2$ وذلك باستخدام أدوات اختبار تعتمد على التوزيع المعتاد المعياري.

ولكنه مع اختبار الفروض الإحصائية المتعلقة بالوسط أو الفرق بين الوسطين والنسبة والفرق بين نسبتي فإنه قد يعنينا أيضاً اختبار درجة تباين ظاهرة أو متغير ما σ^2 أو مقارنة تباين متغير ما بتباين آخر . فعادة ما يكون اهتمام وحدة مراقبة وفحص الإنتاج في مصنع ما - كمصنع للمياه الغازية أو بعض المواد الغذائية السائلة كالزيوت - فقد لا يقتصر الاهتمام فقط على متوسط

حجم المياه الغازية المعبأة في الزجاجات ، ولكن تباين حجم السائل المعبأ والذي قد يختلف من زجاجة لأخرى لأسباب قد يمكن ضبطها أو عدم إمكان التحكم فيها . هذا التفاوت أو التباين لا يقل أهمية لأثره على المستهلك إقبالا أو إعراضا وعلى الجهات أو الأجهزة المعنية بالرقابة على الإنتاج .

(١-١) اختبارات الفروض بشأن تباين مجتمع ما σ^2 :

مثال (١-١) :

إذا رمزنا إلى تباين توزيع حجم المياه الغازية في الزجاجات في مصنع ما بـ σ^2 وتقديرها من العينة التي تسحب دوريا للفحص بـ $\hat{\sigma}^2$ وبفرض أن خطة الفحص الدوري تقضى بأن تكون $\bar{c} > 0.01$ وبفرض أن حجم عينة الفحص $n = 10$ زجاجات وبفرض أن $\bar{c} = 0.04$ فهل يدل ذلك على أن نتيجة الفحص تقدم دليلا كافيا على أن $\hat{\sigma}^2$ تختلف معنويا عن $\sigma^2 = 0.01$ عند مستوى المعنوية 5% ؟ وللوصول إلى قرار في هذا الشأن فيمكن أن نعتمد على توزيع احتمالي هو χ^2 (chi-square) ويعرف في حالتنا هذه كالآتي :

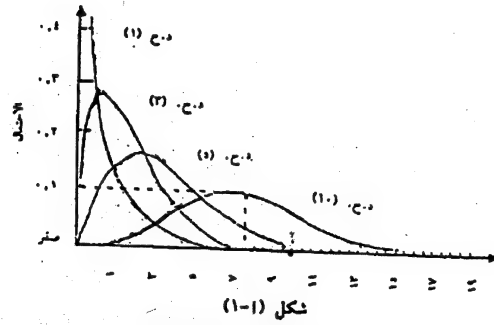
$$\chi^2_{\alpha} = \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \quad (١/١)$$

وهو توزيع المعاينة لهذا المتغير $\hat{\sigma}^2$ عند توافر شروط معينة

توزيع كا^٢ Chi-square distribution :

وهو توزيع إحتمالي اكتشفه عالمي الإحصاء المعروفين فيشر وبيرسن Sir Ronald Fisher & Karl Pearson ونسك في مطلع القرن الماضي . ولهذا التوزيع مؤشر (بارامتر) واحد هو درجات الحرية (ن-١ في مثالنا هذا) وبالتالي فهو ليس توزيعا وحيدا كالتوزيع المعتاد المعياري ، ولكنه مجموعة أو عائلة من التوزيعات تختلف باختلاف درجات الحرية ، كما أنه توزيع موجب مستمر أى يقع بأكمله على يمين المحور الرأسى ، أى أن جميع قيمة موجبة وهو توزيع موجب الالتواء لدرجات الحرية صغيرة العدد ، ويقترّب من التماثل

كلما كبر عدد تلك الدرجات ، ولهذا التوزيع استخدامات عديدة خاصة في الاختبارات اللامعلمية كما سنرى في باب لاحق . والشكل التالي يعرض صورا مختلفة لهذا التوزيع ويقدم جدول رقم (١٠) في نهاية هذا المرجع الاحتمالات أو المساحات تحت منحنى التوزيع لدرجات حرية مختلفة أى أن الذيل الأيمن أو الطرف الأعلى للمنحنى يحدد المساحة $\alpha = P(\chi^2 > \chi^2_{\alpha})$



شكل (١-١) توزيع كاي^٢ لدرجات حرية مختلفة

وتطبيقا لذلك على المثال السابق (١-١) حيث $n = 10$ ، $\sigma^2 = 0.01$ ، $\alpha = 0.04$ فإن الاختبار الإحصائي يتم كالتالى :

الفرض العدمى : $\sigma^2 = 0.01$ والفرض البديل $\sigma^2 > 0.01$

$$\text{أداة الاختبار الإحصائي كاي}^2 = \frac{\sum (x_i - \bar{x})^2}{\sigma^2}$$

لها توزيع كاي^٢ بدرجات حرية $n - 1 = 9$ ويرفض الفرض العدمى إذا كانت القيمة المحسوبة لـ كاي^٢ أكبر من ١٦,٩١٩٠ وهى القيمة الجدولية كاي^٢ ٠,٠٤,٩.

ولكن :

$$\text{كاي}^2 = \frac{(0.04) \times 9}{0.01} = 3.6$$

وبالتالى يقبل الفرض العدمى بأن $\sigma^2 > 0.01$ أو $(\sigma > 0.1)$ باحتمال $\alpha = 0.05$.

ملحوظة : أجرى هذا الاختبار بفرض أن توزيع حجم المياه الغازية فى الزجاجات فى المصنع (أى المجتمع) يقترب من التوزيع الطبيعي أو المعتاد بغض النظر عن حجم العينة .

(٣-١) اختبار الفروض بشأن تباين مجتمعين : $\sigma_1^2 = \sigma_2^2$

عند مقارنة تباين توزيع قراءات تسجل عن مجتمعين مستقلين ، كنتائج اختبارين للقدرات أو الصلاحية الذى يطبق على عينة من المتقدمين لشغل وظيفة مسا أو القراءات التى تسجلها وحدتى قياس (ترمومترين مثلا) فإنه من المهم أن نقارن التباين لأداء المقياسين أو الاختبارين للوصول إلى قرار فى شأن أى من أداة القياس أو الاختبار يمكن استخدامها . وبفرض أن $\sigma_1^2 = \sigma_2^2$ نرمزان إلى تباينى المجتمعين الأول والثانى ، فإنه لاختبار الفرض بتساوى أو عدم تساوى التباينين أى :

$$\text{الفرض العدمى : } 1 = \frac{\sigma_1^2}{\sigma_2^2} \text{ أى } \sigma_1^2 = \sigma_2^2$$

$$\text{الفرض البديل : } 1 \neq \frac{\sigma_1^2}{\sigma_2^2} \text{ أى } \sigma_1^2 \neq \sigma_2^2$$

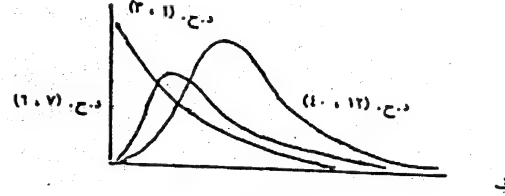
فإن ذلك يتم باستخدام توزيع F-distribution أو توزيع نسبة التباين F Variance ratio حيث :

$$F = \frac{\frac{s_1^2}{n_1}}{\frac{s_2^2}{n_2}} \quad (2/1)$$

توزيع F-Snedecor :

وينسب هذا التوزيع الاحتمالي إلى العالم الإحصائي G. W. Snedecor الذي قدمه تكريماً للعالم الإحصائي فيشر Fisher الذي عرّف هذه الخاصية باسم "نسبة التباين" وقام بإعداد جداول للعلاقة في صورة $Z = \log_e \sqrt{F}$ ثم قام Snedecor بتقديم توزيع نسبة التباين في صورته المجدولة حالياً وهي من الناحية العملية أسهل في الاستخدام خاصة في تحليل التباين (جداول من ٤ إلى ٧ في نهاية هذا المرجع).

وهذا التوزيع ليس توزيعاً وحيداً ولكنه توزيع نسبة متغيرين عشوائيين لكل توزيع كـ^٢ بدرجة حرية $n_1 - 1$ للبسط ، درجة حرية $n_2 - 1$ للمتغير في المقام وبالتالي فهو توزيع يعتمد على معلمتين هما درجات حرية البسط والمقام وبالتالي يختلف باختلاف قيمة هاتين المعلمتين . والشكل التالي يعرض صوراً لهذا التوزيع لبعض درجات الحرية .



شكل (٢-١)

توزيع F لدرجات حرية مختلفة

وتتركز خصائص هذا التوزيع في الآتي :

- ١- أنه توزيع موجب .
- ٢- أنه توزيع موجب الالتواء .

٣- أن منحني التوزيع يقترب من المحور الأفقي بازدياد عدد درجات الحرية دون أن يمس المحور .

ويستخدم هذا التوزيع ، أى توزيع ف لنسبة التباين لاختبار تساوى تباينى مجتمعين مستقلين ، بشرط أن يكون توزيع الظاهرة لكل من المجتمعين توزيع معناد وأن تكون العينتين المسحوبتين من المجتمعين عينات عشوائية مستقلة كل منهما عن الأخرى . كما يستخدم أيضا لاختبار ومقارنة متوسطات عدة مجتمعات وهو الأسلوب المعروف باسم "تحليل التباين" (ANOVA) Analysis Of Variance والذي سنعرض له فى مرحلة لاحقة متى توافرت شروط معينة فى توزيعات المجتمعات المقارنة وأهمها أن تكون توزيعات المجتمعات المقارنة توزيعات معنادة وأن تكون وحدات القياس من فئة القياس بفترة (interval scale) على الأقل وإذا لم يتحقق هذين الشرطين أو ثارت الشكوك حول تحققها فتستخدم أحد الأساليب اللامعلمية بدلا من توزيع ف وهو ما سوف نعرض له مستقبلا .

هذا ونذكر أيضا أنه عند استخدام اختبار ت للفرق بين متوسطين $\mu_1 - \mu_2$ باستخدام أداة الاختبار :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (3/1)$$

حيث ع ت : التباين التجميعي محسوبا من العينتين ، فإنه نفترض تساوى تباين توزيع المجتمعين ، وبالتالي فإنه يتعين استخدام اختبار ف لاختبار تساوى σ_1^2 ، σ_2^2 ،

مثال (١-٣):

بفرض أن شركة لسيارات الليموزين لنقل المسافرين من ميدان التحرير إلى مطار القاهرة ويمكن أن تستخدم سياراتها أحد مسارين - أى طريقين - ولما كان من المهم أن تتحقق الشركة من اتساق وتوافق الوقت المنفق في الوصول إلى المطار باستخدام أي من المسارين فقد سجلت الشركة البيانات التالية:

الطريق (أو المسار)	متوسط الوقت بالدقيقة	الانحراف المعياري بالدقيقة	عدد السيارات
الطريق أ	٥٦	١٢	٧
الطريق ب	٥٨	٥	٨

اختبر الفرض بأن $\sigma_1 = \sigma_2$ عند مستوى المعنوية $\alpha = 10\%$.
ويتم تحقيق ذلك كالآتي:

الفرض العدمي: $\sigma_1 \neq \sigma_2$ أى $\sigma_1^2 / \sigma_2^2 \neq 1$
والفرض البديل $\sigma_1 \neq \sigma_2$ أى $\sigma_1^2 / \sigma_2^2 \neq 1$ أى أنه اختبار في اتجاهين ، وبالتالي فإن مستوى المعنوية يصبح $\alpha/2 = 5\%$ وهى مساحة الطرف الأعلى (أو الأيمن) من التوزيع وهى القيم الواردة بجدول توزيع ف.

$$\text{أداة الاختبار : } F = \frac{s_1^2}{s_2^2}$$

وبرفض الفرض العدمي إذا كانت F^* (أي المحسوبة من البيانات) أكبر

$$\text{من } F_{\alpha/2, n_1-1, n_2-1} = 3.87$$

$$\text{ولكن } F^* = \frac{12^2}{5^2} = 5.76 < 3.87 \text{ ويرفض الفرض العدمي .}$$

ولكنه من المعتاد أن نحسب $F^* = \frac{s_1^2}{s_2^2}$ ، حيث s_1^2 هى الأكبر قيمة (أى $s_1^2 > s_2^2$) وفى هذه الحالة فإنه إذا كان الفرض العدمي هو :

$\sigma^2 \geq \sigma^2_0$ والفرض البديل هو $\sigma^2 < \sigma^2_0$ وعند مستوى المعنوية $\alpha = 10\%$ وهو اختبار في اتجاه واحد فيرفض الفرض العدمي إذا كانت

$$F = \frac{s^2_E}{s^2_C} = \frac{12}{5} = 2.4 < 5.76 = F_{0.10, 7, 16} = 5.83$$

وبالتالي فإن نتيجة الاختبار تؤيد قبول الفرض البديل أي أن تباين توزيع الوقت المنفق في قطع المسافة بأى من المسارين يختلف معنوياً عن $\alpha = 10\%$.

(٣) الاختبارات الإحصائية المتعلقة بالفروق بين متوسطين والفروق بين عدة متوسطات :

درسنا في مرحلة سابقة الاختبارات الإحصائية للفروق بين متوسطين $(\mu_1 - \mu_2)$ وذلك اعتماداً على توزيع العينات للفروق $(\bar{s}_1 - \bar{s}_2)$ سواء باستخدام التوزيع المعتاد المعياري أو توزيع ت كل بفرض تحقق شروط معينة. ولكن كثيراً ما تدعوا الحاجة إلى اختبار الفروق بين عدة متوسطات ، فقد تقوم مؤسسة إنتاجية ما إلى تقديم أنواع مختلفة من الحوافز النقدية أو العينية ولتعيين أى من تلك الحوافز أكثر أثراً على الإنتاجية ، فتقوم تلك المؤسسة بتطبيق أنواع الحوافز المختلفة على عينات عشوائية مستقلة من العمال للوصول إلى قرار في هذا الشأن . فقد تكون الحوافز نقدية بنسبة ثابتة من الأجر أو المرتب بالإضافة إلى تلك نسبة من الأرباح ، وقد تكون في صورة عينية أو أجازات بأجر ... الخ . وفي حالات كهذه وبفرض أننا نقارن بين أثر خمسة أنواع من الحوافز فسوف تكون هناك ١٠ اختبارات (F_{ij}) يتعين إجراؤها وهي تتمثل في اختبار الفروق بين كل متوسطين على حدة وهي: $(\mu_1 - \mu_2)$ ، $(\mu_1 - \mu_3)$ ، $(\mu_1 - \mu_4)$ ، $(\mu_1 - \mu_5)$ ، $(\mu_2 - \mu_3)$ ، $(\mu_2 - \mu_4)$ ، $(\mu_2 - \mu_5)$ ، $(\mu_3 - \mu_4)$ ، $(\mu_3 - \mu_5)$ ، $(\mu_4 - \mu_5)$.

ومن الواضح أن إجراء اختبارات كهذه يعاب عليها الصعوبة العملية التي تنشأ عن تعدد الاختبارات ، علاوة على أن تعدد الاختبارات يؤدي إلى زيادة حجم المنطقة الحرجة عن المقرر أصلاً قبل بدء الاختبارات ، وهو يعني الحكم برفض بعض الفروض التي كان يجب أن تقبل لو لم تتكرر المقارنات. وقد تبين أنه بتكرار اختبار ت عند مستوى المعنوية ٥% للفرق بين متوسطين عدة مرات وكانت هذه الاختبارات مستقلة ، فإن احتمال الحكم بمعنوية أحد هذه الفروق على الأقل يتجاوز الـ ٥% ليصل إلى ٢٣% إذا أجرى هذا الاختبار ٥ مرات ويزداد عن ذلك كلما ازداد عدد مرات تكرار الاختبار.

ولقد دعى ذلك الإحصائي المعروف سير رونالد فيشر (١٩٣٨) Sir Ronald Fisher إلى تقديم أسلوب تحليل التباين (Analysis of Variance) لتحليل الفروق بين المتوسطات حين تتعدد المقارنات . وسوف نعرض في الفقرات التالية لمفهوم عام لهذا الأسلوب ، ثم ننقل بعد ذلك إلى استخدام هذا الأسلوب في تحليل نتائج التجارب التي تصمم بطريقة أو بأخرى لاختبار أثر عامل أو عاملين أو ثلاثة عوامل.

(٣) تحليل التباين (Analysis of Variance) :

سوف نستخدم المثال التالي ببيانات فرضية في تقديم موضوع تحليل التباين .

مثال (١-٣) :

يفرض أن البيانات التالية تمثل عدد الدقائق التي استغرقت في كتابة صفحة من تقرير ما على ثلاثة أنواع مختلفة من الآلات الكاتبة (أ ، ب ، ج) وقد اختبرت عينة عشوائية من ٩ من الناسخين على الآلة الكاتبة ، وروعى في اختيارهم التقارب والتجانس في القدرة على الكتابة على الآلة الكاتبة ثم توزيعهم عشوائياً وبالتساوى على الآلات الثلاث ، بحيث خصص منهم ٣ للكتابة على كل من الأنواع الثلاثة من الآلات:

نوع الآلة	عدد الدقائق التي استغرقت في كتابة الصفحة سر	مجموع س ر	مجموع (س ر - س ن)²
أ	٤	١٨	١٤
ب	٦	١٨	٢
ج	٨	٣٦	٣٢
		٧٢	٤٨

ويمكننا الوصول إلى تقديرين للتباين لكل عينتين ممكنتين وعددها ٣ أزواج ، أى يمكن إجراء ثلاث مقارنات وذلك على النحو التالي:

التقدير الأول : التباين التجميعى Pooled Variance

$$ع^2 = \frac{1}{n_1 + n_2 - 2} [\text{مجموع (س - س}_1\text{)}^2 + \text{مجموع (س - س}_2\text{)}^2] \quad (٤/١)$$

حيث ترمز ١ ، ٢ فى دليل ن ، س إلى العينتين الأولى والثانية (مثلا).

التقدير الثانى: التباين الكلى Total Variance

$$ع^2 = \frac{1}{n - 1} \text{مجموع ر (س ر - س ن)}^2 \quad (٥/١)$$

ويعرف هذا التقدير باسم التباين الكلى، حيث $n = n_1 + n_2 =$ مجموع مفردات العينتين ، س .. الوسط الحسابى العام محسوبا من جميع مفردات العينتين معا.

وتطبقا للتقديرين السابقين وبصرف النظر عن المقام بالاختصار أى على البسط عند حساب مجموع المربعات . فسوف نجد الاتى محسوبا من البيانات السابقة .

بسط التباين الكلى	بسط التباين التجميعى
مجـ (مكرر - م...) ^٢ ن - م + ن وهو مجموع المربعات الكلى	مجـ ر [مجـ ر (س - م...) ^٢] حيث ر = ١، ٢، ٣ ل = أ أو ب أو جـ وهو مجموع المربعات داخل العينات
	الأكتين أ ب :
م... - م... = م... - م... مجموع المربعات = ١٦	مجموع المربعات = ١٦ = ٢ + ١٤
	الأكتين أ ، جـ :
م... - م... = م... - م... مجموع المربعات = ١٠٠	مجموع المربعات = ٤٦ = ٢٣ + ١٤
	الأكتين ب ، جـ :
م... - م... = م... - م... مجموع المربعات = ٨٨	مجموع المربعات = ٣٤ = ٣٢ + ٢

ويتضح من النتائج السابقة أن :

مجـ ر (س رز - م...)^٢ - مجـ ر [مجـ ر (س ر - م...)^٢]
= مجموع المربعات الكلى - مجموع المربعات داخل العينات
بالنسبة للأكتين أ ، ب = ١٦ - ١٦ = صفر وكانت م... = م...
وبالنسبة للأكتين أ ، جـ = ٤٦ - ١٠٠ = ٥٤ ، م... ≠ م...
وكذلك في حالة الأكتين ب ، جـ = ٣٤ - ٨٨ = ٥٤ ، م... ≠ م...
وأخيرا في حالة الآلات الثلاث أ ، ب ، جـ ، حيث م... = ٨ فإن :
مجـ ر (س رز - م...)^٢ = ١٢٠ - ٤٨ = ٧٢ ، م... = م... ≠ م...

ومما سبق يتبين لنا أنه عندما كان $\bar{S}_M = \bar{S}_N, \bar{M} \neq \bar{N}$ ، فإن :

مجموع المربعات الكلى = مجموع المربعات داخل العينات

فى حين أنه عندما كان $\bar{S}_M \neq \bar{S}_N, \bar{M} \neq \bar{N}$ فإن :

مجموع المربعات الكلى - مجموع المربعات داخل العينات < صفر

= ٥٤ فى حالة العينتين لكل من (أ، جـ)، (ب، جـ)

= ٧٢ فى حالة العينات الثلاث (أ، ب، جـ)

والطرف الأيسر يشير إلى مجموع المربعات الذي ينشأ بسبب اختلاف المتوسطات ، أي اختلاف المتوسطات عن المتوسط العام ، ويعرف عادة باسم مجموع المربعات بين العينات ، وهذه هي الفكرة من تحليل التباين. وبالتالي فإنه يمكن تعريف أسلوب تحليل التباين بأنه " أسلوب إحصائي يعتمد فى اختبار أثر أنواع مختلفة من عامل (أو أكثر) على اختلاف المتوسطات التي تعبر عن اختلاف الأداء باختلاف نوع أو مستوى العامل المستخدم وذلك بتجزئة مجموع المربعات الكلى إلى مركبات يعزى إحداها إلى الاختلاف بين المتوسط محسوبا لكل نوع عن المتوسط العام والمركبة الأخرى إلى الاختلاف داخل كل عينة أي اختلاف مفردات كل عينة عن متوسطها ويعرف باسم البواقي (Residual) أي أنه فى حالة التحليل لعامل واحد سنصل إلى النتيجة التالية:

مجموع المربعات الكلى = مجموعات المربعات بين العينات (الأنواع) +

مجموع المربعات داخل العينات (البواقي).

وبالتالى فإنه يمكن تلخيص وتحليل النتائج الكلية للمثال السابق باستخدام

أسلوب تحليل التباين على النحو التالي:

(١) مجموع المربعات الكلى الذي يعبر عن التباين الكلى أي أنه يعزى إلى

اختلاف الوقت المنفق فى كتابة صفحة التقرير بكل من كاتبى الآلة

الكاتبة الـ (٩) أي س زر حيث يرمز لـ = أ أو ب أو جـ إلى نوع

الآلة الكاتبة ، ر : ١ ، ٢ ، ٣ إلى الناسخين الثلاث بكل آلة كاتبة عن المتوسط العام لوقت الكتابة أي $\bar{S} \dots$ ، أي أن :
مجموع (س.ر - $\bar{S} \dots$) = ١٢٠ - ٢ ويعرف باسم مجموع المربعات الكلي أو م.م.ك.

(٢) مجموع المربعات الذي يعزى إلى اختلاف متوسط الوقت الذي استغرقه كل ناسخ باستخدام الآلة التي خصصت له من الآلات الثلاث أي $\bar{S} \dots$ ر. عن المتوسط العام $\bar{S} \dots$ ويساوي

$$= \text{مجموع} (\bar{S} \dots - \bar{S} \dots)^2 = 3(8-6)^2 + 3(8-6)^2 + 3(8-12)^2 = 72$$

ويعرف هذا المجموع باسم مجموع المربعات بين أنواع الآلات (م.م.ب)

(٣) مجموع المربعات الذي يعزى إلى اختلاف الوقت الذي استغرقه كل ناسخ من الناسخين الثلاث على كل آلة من الأنواع الثلاث عن متوسط الوقت المستغرق في الكتابة على الآلة الواحدة من كل نوع ، أي :

$$\text{مجموع} (\text{س.ر} - \bar{S} \dots)^2 = 14 + 2 + 32 = 48 \text{ وهو يساوي الفرق بين}$$

م.م.ك ، م.م.ب أي يساوي ١٢٠ - ٧٢ = ٤٨ ولذلك فهو يعرف باسم مجموع

المربعات المتبقية أو البواقي (Residual) ويرمز له بالرمز (م.م.خ) ويشير إلى التباين العشوائي بين الناسخين على كل آلة أو التباين داخل العينات الثلاث

(Within). أي أنه يمكن صياغة هذه العلاقة كالآتي :

مجموع المركبات الكلي (م.م.ك) = مجموع المربعات بين أنواع الآلات (م.م.ب)

+ مجموع المركبات داخل العينات أو البواقي (م.م.خ) (٦/١)

أي :

$$S.S. \text{ Total} = S.S. \text{ Between} + S.S. \text{ Residual}$$

$$\text{ورقمياً من المثال : } 120 = 72 + 48$$

ويمكن تصوير هذه النتائج في جدول يعرف باسم جدول تحليل التباين

(ANOVA) وهو على الصورة التالية:

جدول تحليل التباين

المصدر	مجموع المربعات (م.م)	درجات الحرية (د.ح.)	متوسط م.م. (م.م.م)	نسبة التباين (ف)
بين أنواع الآلات (أو بين العينات)	$\sum (x_i - \bar{x})^2$	$(I - 1)$	$\frac{\sum (x_i - \bar{x})^2}{(I - 1)}$	$\frac{\sum (x_i - \bar{x})^2}{\sum (x_{ij} - \bar{x}_{..})^2}$
داخل العينات	$\sum \sum (x_{ij} - \bar{x}_i)^2$	$I(J - 1)$	$\frac{\sum \sum (x_{ij} - \bar{x}_i)^2}{I(J - 1)}$	
كلى	$\sum \sum (x_{ij} - \bar{x}_{..})^2$	$IJ - 1$		

وتطبقا على المثال السابق:

المصدر	مجموع المربعات (م.م)	درجات الحرية (د.ح.)	متوسط م.م. (م.م.م)	نسبة التباين (ف)
بين العينات	72	2	36	$8 \div 36 = 0.22$
داخل العينات	48	6	8	
كلى	120	8		

(** يمكن أن تحسب القيم المناظرة بالفرق)

وتقارن قيمة ف* (أى المحسوبة) بقيمة ف لتوزيع ف عند درجات الحرية 2 ، 6 ، وعند مستوى المعنوية α المحدد سلفا ويتخذ على أساس مقارنة ف* بـ ف الجدولية يكون القرار بقبول أو رفض الفرض العدمى .
فمثلاً عند $\alpha = 5\%$ وكان الفرض العدمى هو أن متوسط سرعة الأداء على الآلات الثلاث متساو أي $\mu_1 = \mu_2 = \mu_3$ ، مقابل الفرض البديل أن الأداء مختلف باستخدام نوعين على الأقل من الآلات الكتابية المستخدمة وعند القيمة الجدولية ف 0.05 ، 6 ، 2 = 5.99 فإن ف* = 0.22 > 5.99 لذلك يقبل الفرض العدمى بعدم اختلاف سرعة الأداء باستخدام أنواع الآلات الثلاث.

وفي الحقيقة فإن الاختبارات تكون للفروق بين المتوسطات $\mu_1 - \mu_2$ ،
 $\mu_2 - \mu_3$ حيث لا يوجد مبرر لاختبار الفرق من $\mu_1 - \mu_3$ بالتساوي
 المتوسطين \bar{S}_1 ، \bar{S}_2 .

ولاختبار الفرض العدمي بأن $\mu_1 = \mu_2$ أى $\mu_1 - \mu_2 = 0$ ، صفر ، مقابل الفرض
 البديل بأن $\mu_1 \neq \mu_2$ وبإستخدام اختبار ت وهو:

$$T = \frac{(\bar{S}_1 - \bar{S}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma^2}}$$

$$T = \frac{12-6}{\sqrt{\left(\frac{2}{3}\right) 46}} = \frac{6}{5.038} = 1.193$$

وكذلك لاختبار الفرض العدمي $\mu_1 - \mu_2 = 0$ صفر مقابل $\mu_1 - \mu_2 \neq 0$

$$T = \frac{2-6}{\sqrt{\left(\frac{2}{3}\right) 46}} = \frac{-4}{5.038} = -0.794$$

وبمقارنة ت فى الحالتين بقيمة ت الجدولية عند $\alpha = 0.05$ ، ٤ درجة
 حرية ، نجدها ± 2.776 وكلاهما أصغر من قيمة ت الجدولية ، لذلك يقبل
 الفرض العدمي فى الحالتين.

وهذه النتائج تتفق مع ما انتهى إليه اختبار الفرض الإحصائي العدمي
 بأن $\mu_1 = \mu_2 = \mu_3$ والذي استخدم فيه اختبار نسبة التباين ف فى الفقرة
 السابقة.

(٤) ملاحظات ختامية :

(١) يفترض لصحة استخدام أسلوب تحليل التباين لمقارنة متوسطات عينتين أو أكثر ، أن العينات قد سحبت من مجتمعات متساوية في الأداء مع توافر شروط ثلاث هي :

- (١/١) أن يكون توزيع المتغير موضوع القياس في المجتمعات التي سحبت منها العينات لها التوزيع المعتاد .
- (٢/١) وأن σ^2 واحدة لتلك المجتمعات .
- (٣/١) أن تكون العينات العشوائية مستقلة .

(٢) يمكن أن يحسب مجموع المربعات الكلى وكذلك مجموع المربعات بين أنواع الآلات وداخل العينات ، بطرق جبرية مبسطة غير ما عرض في الفقرة السابقة ، وهو ما سوف نتناوله بالتفصيل في مرحلة لاحقة في تحليل نتائج التجارب .

(٣) يستخدم أسلوب تحليل التباين في تحليل الانحدار البسيط والمتعدد وذلك بتجزئه مجموع المربعات الكلى إلى مركبتين (أو أكثر) لتحديد واختبار أثر العلاقة البسيطة أو المتعددة للمتغير (أو المتغيرات المستقلة) على المتغير التابع أو غير المستقل ، وهو ما سوف يعرضه هذا المرجع عند دراسة تحليل الانحدار . وفي الحقيقة فإن أسلوب تحليل التباين وتحليل الانحدار ليس إلا تطبيق لأسلوب المربعات الصغرى Method of Least Squares .

وسوف نتعرض في بقية هذا الفصل إلى مدخل في تصميم التجارب Design of Experiments وسوف نبدأه بمعالجة تصميم وتحليل نتائج التجربة في اتجاه واحد ، وهو المعروف باسم التصميم كامل العشوائية ، ثم ننقل إلى التصميم في اتجاهين أو أكثر بقدر ما يسمح به مستوى العرض

- لهذا المرجع ، ويمكن للقارئ متابعة هذا الموضوع بالرجوع إلى بعض المراجع الأكثر تخصصاً الواردة في نهاية هذا الكتاب.
- (٤) وأيا كان النموذج الذي يتبع في تصميم وتحليل نتائج التجربة ، فالهدف لا يخرج عن تحقيق واحد أو أكثر من الأهداف التالية:
- (١/٤) اختبار أثر الأنواع أو المستويات المختلفة لعامل واحد ومعنوية تأثيرها على وحدات التجربة.
- (٢/٤) اختبار متوسط الأداء لنوع ما أو لمستوى معين للعامل موضوع التجربة والدراسة أو اختبار الفرق بين متوسطي (أو متوسطات) الأداء للأنواع أو المستويات المختلفة لذلك العامل وكذا تعيين فترات ثقة لمتوسطات الأداء.
- (٣/٤) قياس وتقدير الكفاءة النسبية لنموذج ما مقارنة بنماذج أخرى يمكن استخدامها.
- (٤/٤) تقدير مركبات التباين للمراحل المختلفة للتجربة ومكوناتها (بين العوامل وداخل العينات وداخل العينات الفرعية ... إلخ) وهو ما يمكن الاستفادة منه في إعادة تخطيط النموذج الذي يحقق كفاءة أعلى أو بخفض من نفقات التجربة إذا ما أعيدت في المستقبل.

ثانياً : تصميم التجارب Design of Experiments

وسوف نقدم فيما يلي بعض التعاريف السائدة في هذا المجال:

تعريف:

(١-١) التجربة Experiment:

انتهت الفقرات السابقة إلى أنه حين تتعدد المقارنات بسبب تعدد المتغيرات موضوع الدراسة أو عدم استقلال بعضها عن البعض ، فإنه تقوم الحاجة إلى تجربة تلك العوامل أو بعضها لقياس أثرها . وعادة ما تقوم التجربة: وهى استقصاء مخطط بهدف الحصول على معلومات مدققة عن متغير ما يعرف فى مجال تحليل الانحدار باسم المتغير التابع أو المتغير غير المستقل (Dependent Variable) كما يعرف أيضا باسم متغير الاستجابة (Response Variable) فى مجال تصميم التجارب ، وقياس أثر واحد أو أكثر من المتغيرات المستقلة عليه وتعرف باسم العوامل (Factors) أو المعالجات (Treatment) أو القطاعات (Blocks) أو التفاعلات (Interactions) فى مجال تصميم التجارب.

ولقد سادت أعمال الإحصائي الكبير R.A.Fisher (١٩١٨ - ١٩٣٨) مجال تصميم التجارب حين كان مسئولاً عن الإحصاء فى محطة التجارب الزراعية ، الأمر الذى أدى إلى أن يسود اسم " المعالجات " و "القطاعات" و "التفاعلات" و "القطع" وهى مستمدة من المجال الزراعى - لغة تصميم التجارب ثم امتد الأخذ بهذه المسميات فى كافة مجالات تصميم التجارب فى العلوم الإنسانية والطبية والصنيدلية ... إلخ ، وسوف نورد بعض المصطلحات السائدة فى هذا المجال قبل الدخول فى تناول التصميم والتحليل الإحصائي للتجربة ونتائجها.

(٣-١) المعالجات Treatments :

وتطلق على مستويات العامل موضوع الدراسة ، وحين يكون العامل وحيداً كالأجر في صناعة أو عمل ما بحسب النوع (ذكور وإناث) ، فإن هذا العامل يكون ذا مستويين . وحين تشمل التجربة عاملين أو أكثر فإنه يمكن قياس الأثر الأساسي لكل عامل بالإضافة إلى الأثر (أو الآثار) المشتركة لمستويات العوامل والتي تعرف باسم تفاعلات.

(٣-١) وحدة التجربة Experimental Unit :

وهي الوحدة التي تسجل عليها استجابة العامل أو العوامل موضوع الدراسة في التجربة ، وسيكون مجال دراستنا للتجارب المخططة " Designed Experiment " أي التي يتم فيها ضبط والتحكم في مواصفات المعالجات وكذلك طريقة تخصيص أو توزيع وحدات التجربة على المعالجات . فحين يكون موضوع الدراسة هو قياس أثر اختلاف وسائل الإعلان على حجم أو قيمة المبيعات من سلعة ما بين وسائل مرئية (الإعلانات بالصحف والستلفزيون مثلاً) أو توزيع عينات مجانية عن المنتج لفترة ما ، أو وسائل مسموعة كالإذاعة ، فإن طريقة الإعلان عن المنتج هي المعالجات ووحدة التجربة هي قيمة المتفق على الإعلان في أي من هذه الوسائل .

(٤-١) خطأ التجربة Experimental Error :

وبشير إلى الاختلاف بين الوحدات للتجريبية التي تعامل بمعالجة واحدة بسبب الاختلاف في طرق المعالجة أو عدم التجانس بين وحدات التجربة أو القصور في إجراء التجربة.

التكرار وأهميته Replicates : أي تكرار التجربة على أكثر من وحدة تجريبية ، وذلك يمكننا من قياس أو تقدير خطأ التجربة وكذلك الارتفاع بدرجة دقة التجربة بتصغير قيمة ع' وكذلك لإمكان التعميم من نتائج التجربة بعد إجراء الاختبارات الإحصائية المناسبة.

(0-1) العشوائية Randomization :

وهي ضرورية لتحقيق التجانس بين وحدات التجربة ولتأمين الباحث من الوقوع في خطأ التحيز.

تحليل التباين ANOVA وهو السبيل إلى تحليل نتائج التجربة واختبار الفروض موضوع البحث ، والتعميم لنتائج التجربة من العينة إلى المجتمع الأكبر والأوسع الذي سحبت منه العينة.

وسوف ننقل إلى معالجة موضوع تصميم التجارب بتقديم بعض النماذج التي تستخدم في هذا المجال ، وسوف نقصر على البعض منها وعلى الأخص:

- التصميم كامل العشوائية Completely Randomized Design
- تصميم القطاعات الكاملة العشوائية Complete Randomized Blocks Design
- المربع اللاتيني .
- التحليل العاملي .

وسوف نقدم في بقية هذا الفصل التصميم أو النموذج كامل العشوائية ثم نقدم بقية التصميمات في الفصل الثاني ، ويمكن للقارئ أن يرجع إلى مراجع أخرى أكثر تخصصاً ذكر البعض منها في قائمة المراجع في نهاية هذا المرجع وذلك للتعرف على التصميمات الأخرى.

ثالثاً : التصميم كامل العشوائية

التصنيف في اتجاه واحد

Completely Randomized Design One way classification

(١) استخدام عدد متساو من وحدات التجربة في كل معاملة

يستهدف هذا التصميم دراسة أثر أنواع مختلفة لعامل واحد على وحدات تجربة عشوائية ، أي بحيث توزع مفردات التجربة بطريقة عشوائية على الأنواع المختلفة للعامل موضوع الدراسة ، ولذلك يعرف هذا النموذج باسم التحليل لعامل واحد (One or Single Factor Analysis) كما يعرف أيضاً باسم التصنيف أو التحليل في اتجاه واحد (One Way Classification) . فإذا كانت الدراسة تستهدف - مثلاً - اختبار أثر أنواع مختلفة من الحوافز على إنتاج عمال وحدة إنتاجية ، وبفرض أن أنواع الحوافز ل = ٣ (نقدية ، عينية ، أخرى) وكانت وحدات التجربة ١٥ عاملاً ، فإنه يمكن توزيع العمال الـ ١٥ على الأنواع الثلاث من الحوافز بطريقة عشوائية ، وقد يتساوى عدد العمال الذين يخصصون لتجربة كل من الأنواع الثلاث أي يختار $r = ٥$ من العمال عشوائياً لتجربة أثر كل نوع من الأنواع الثلاث من الحوافز ، وأن لم يكن ذلك ضرورياً رغم أن تخصيص عدد متساو لكل نوع سوف يؤدي إلى خفض في الجهد الحسابي.

(١-١) النموذج الرياضي:

وسواء تساوى عدد مفردات التجربة لكل نوع أو لم يتساوى ، فإن نموذج هذه التجربة سيكون كالتالي:

(٧/١)

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

حيث ل = ١ ، ٢ ، ٠٠٠ ترمز إلى الأنواع المختلفة للعامل موضوع الاختبار

ويطلق عليها عادة اسم معالجات (Treatments) .
 $r = 1, 2, \dots, n$ عدد وحدات التجربة في كل معالجة .
 μ = المتوسط العام وهو ثابت مجهول القيمة
 m = وترمز إلى الأثر المضاف للمعالجة ل.
 χ رد = الخطأ العشوائى ، ويشير إلى أثر العوامل العشوائية أو التى لا تخضع للقياس أو تكون تحت السيطرة فى التجربة.
 s رد : وتشير إلى ناتج التجربة فى وحدة التجربة أو المفردة r من المعالجة ل.

ويفترض فى بناء هذا النموذج الآتى:

- ١- أن ناتج التجربة s ينشأ عن آثار مضافة (Additive) ، هى أثر المعالجة ل على وحدة التجربة r ممثلاً فى المتوسط العام μ مضافاً إليه أثر المعالجة ل علاوة على تفاوت أو تباين عشوائى .
- ٢- m ر تعبر عن الأثر المضاف الذى ينشأ بسبب المعالجة ل .
- ٣- إن الأخطاء χ ر تعبر عن أثر العوامل العشوائية وتشمل فى ذلك العوامل التى لا تخضع لسيطرة من يقوم بالتجربة ، ومنها التباين الداخلى بين مفردات التجربة ، كما يشمل أثر العوامل الطارئة التى قد لا تتكرر أو تستكرر دون انتظام . وسوف يفترض أن لهذه الأخطاء توزيع معتاد توقعه الصفر وتباينه σ^2 وإنها مستقلة عن m ر وهو شرط ضرورى لإجراء الاختبارات والتقدير بفترة ثقة .
- ٤- إن القراءات أو القياسات التى تسجل فى التجربة من النوع الكمية لمتغير مستمر وليست من القياسات التصنيفية أو الترتيبية .
 ويتعين لتنفيذ تجربة كهذه أن يكون توزيع مفردات التجربة عشوائياً على أنواع العامل أو المعالجات ، وبذلك تتاح فرص متساوية لكل معالجة ، كما أن العشوائية توفر استقلال توزيع الأخطاء وألا تتعرض للارتباط الداخلى أو

(٢-١) النموذج الحسابي والتحليل :

$$+ \text{مجبور} (س.ر - \bar{س.ر}) = \text{مجبور} (س.ر - \bar{س.ر}) + \text{مجبور} (س.ر - \bar{س.ر})$$

حيث \bar{S}_r : الوسط الحسابي للمشاهدات في المعالجة ل = ١ ، ٢ ، ..

$$= \text{مـجـ رـ} \frac{\text{سـ رـ}}{\text{ن}}$$

سرر : المشاهدة الرائية في المعالجة ل

يعزى إلى الاختلاف بين المعالجات فى الأثر) + مجموع المربعات

والعلاقة (٨/١) عادة تكتب على النحو التالي كصيغة حسابية :

$$(۱۰/۱) \quad (م) \frac{r}{nd} + (م) ر - مج د$$

حيث م : المجموع الكلى لقيم سر = مجموع المشاهدات
 ن : عدد المفردات الكلى فى التجربة .
 م : مجموع المشاهدات فى المعالجة ل
 ن: عدد المفردات أو المشاهدات للمعالجة ل = ر فى حالة تساوى عدد
 الوحدات التجريبية فى كل معالجة .
 وعادة ما يحسب الجانب الأيمن من (١٠/١) ثم الحد الأول من الجانب
 الأيسر ، أما الحد الأخير فيحسب بالفرق أى أن :
 مجموع المربعات الكلى - مجموع المربعات بين المعالجات = مجموع
 المربعات المتبقى (أو الذى يعزى إلى التباين العشوائى) .
 ويتم تصوير هذه النتائج فى جدول خاص يعرف باسم جدول تحليل
 التباين (Analysis Of Variance) ويختصر عادة بالرمز ANOVA على
 النحو التالى:

تحليل التباين

المصدر Source	مجموع المربعات (م.م) Sum of Squares (s.s)	درجات الحرية (ج.د) Degrees of Freedom (D.F.)	متوسط المربعات (م.م) Mean Squares (M.S.)	نسبة التباين (ف) Variance Ratio (V.R.)
بين المعالجات Between Treatments	(١) مجموع $\frac{م^2}{ن}$	(٢) (١ - ل)	م.م.م.م = (١) ÷ (٢)	
داخل المعالجات (البواقي أو الخطأ العشوائى) Within Treatments or Residual or Error	(٣) بالفرق	(٤) (ن - ل) أو (بالفرق)	م.م.م.م = (٣) ÷ (٤)	م.م.م.م = م.م.م.م
الكلى Total	مجموع من أ و ب - ن	(ن - ١)		

الفرض العدمى : $\mu_1 = \mu_2 = \dots = \mu_r$
 والفرض البديل : $\mu_1 \neq \mu_2 \neq \dots \neq \mu_r$ لأى معالجتين على الأقل .

ولاختبار الفرض السابق ، فإن هذا الفرض يرفض عند مستوى المعنوية α إذا كانت نسبة التباين من العمود الأخير من جدول تحليل التباين أى ف * =

$$\frac{م.م.م}{م.م.م} < ف (ل-١، ن-١، و) \alpha \text{ (ويقبل فرض العدم فيما عدا ذلك .}$$

مثال (١-٤) :

استخدم أسلوب تحليل التباين (نموذج التحليل فى اتجاه واحد) فى دراسة اختلاف إنتاجية ثلاثة أنواع من الآلات ، وذلك بتوزيع أفراد عينة عشوائية مكونة من ١٥ عاملاً عشوائياً على كل من الأنواع الثلاثة من الآلات ، حيث تمثل القراءات المشاهدات المسجلة ، أى عدد الوحدات المنتجة باستخدام هذه الآلات الثلاث فى نهاية اليوم .

الآلة	عدد الوحدات المنتجة س						س.ر.
أ	٤٧	٥٣	٤٩	٥٠	٤٦	٢٤٥	٤٩
ب	٥٥	٥٤	٥٨	٦١	٥٢	٢٨٠	٥٦
ج	٥٤	٥٠	٥١	٥١	٤٩	٢٥٥	٥١

$$\text{م.ر.ر.س}^2 = ٤٠٧٨٤ ، \text{م} = ٧٨٠ ، \text{س.ر.} = ٥٢$$

وبتحليل نتائج هذه التجربة باستخدام أسلوب تحليل التباين نجد أن :

$$\text{مجموع المربعات الكلى} = \text{م.ر.ر.س}^2 - \frac{\text{م}^2}{\text{ن}}$$

$$= ٤٠٧٨٤ - \frac{٧٨٠^2}{١٥} = ٤٠٧٨٤ - ٣٩٠٤٠ = ١٧٤٤ = \text{م.م.ك}$$

مجموع المربعات بين الآلات (أى الذى يعزى إلى اختلاف أثر أنواع الآلات

$$\text{على الإنتاج} = \text{م.ر.ر.س}^2 - \frac{\text{م}^2}{\text{ن}} \text{ حيث } ١ = \text{ن} - ٢ = \text{ن} - ٣ = \text{ر} = ٥$$

$$= \frac{١}{٥} [٢٤٥^2 + ٢٨٠^2 + ٢٥٥^2] - \frac{٧٨٠^2}{١٥} = ١٣٠ = \text{م.م.ب}$$

وهو في نفس الوقت - ر مجر (س.ج. - س.ج. ٠٠) $\sum (س.ج. - س.ج. ٠٠)^2$
 $\sum (٥٢ - ٤٩)^2 + \sum (٥٢ - ٥٦)^2 + \sum (٥٢ - ٥١)^2 = ١٣٠$
 $\sum (١ + ١٦ + ٩) = ١٣٠$
 = مجموع المربعات للفرق بين متوسط الإنتاج لكل نوع من أنواع الآلات والمتوسط العام .

مجموع المربعات داخل المعالجات (أى مجموع المربعات للخطأ العشوائى) = م.م.ك - م.م.ب

= ٢٢٤ - ١٣٠ = ٩٤ = مجموع المربعات المتبقى
 وهو في نفس الوقت = مجر [مجر (س.ج. - س.ج. ٠٠) $\sum (س.ج. - س.ج. ٠٠)^2$]
 $\sum (٤٩ - ٥٠)^2 + \sum (٤٩ - ٤٩)^2 + \sum (٤٩ - ٥٣)^2 + \sum (٤٩ - ٤٧)^2 =$
 $\sum (٥٦ - ٥٨)^2 + \sum (٥٦ - ٥٤)^2 + \sum (٥٦ - ٥٥)^2 + \sum (٤٩ - ٤٦)^2 +$
 $\sum (٥٦ - ٦١)^2 + \sum (٥٦ - ٥٢)^2 + \sum (٥١ - ٥٤)^2 + \sum (٥١ - ٥٠)^2 +$
 $\sum (٥١ - ٥١)^2 + \sum (٥١ - ٤٩)^2 = ٩٤$
 = مجموع المربعات للفرق بين المشاهدات فى كل معالجة والوسط الحسابى للمعالجة .

ويكون جدول تحليل التباين كالاتى :

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بين الآلات	١٣٠	٢	٦٥	
البواقي (أو الخطأ العشوائى)	٩٤	١٢	٧,٨٣	٨,٣٠ -
الكلى	٢٢٤	١٤	-	-

الفرض العدمى = م١ = م٢ = م٣ = صفر
 والفرض البديل هو أن متوسط أثر المعالجات يختلف لمعالجتين على الأقل .

ويرفض الفرض العدمي حيث :

$$\text{نسبة التباين ف} = 8,30 < \text{ف} 80,12,2 = 3,89$$

ملحوظة :

كان يمكن تخفيض الجهد الحسابي بشكل واضح باستبدال قيم س رر بالقسيم س ر - و = حر حيث و : وسط فرض مناسب (٥٠ مثلا في حالتنا هذه) وذلك تطبيقاً للقاعدة المعروفة بأن التباين ع^٢ لا تختلف قيمته العددية سواء حسبت من القراءات الحقيقية أو من القراءات المختزلة بطرح (أو إضافة) مقدار ثابت وهي من أسميناه من قبل سواء بالطريقة المباشرة أو بطريقة الفروق (أو الانحرافات) البسيطة .

(٣) التصميم كامل العشوائية (عدد وحدات التجربة لكل معاملة غير

متساو) :

ولا يختلف الأمر سواء بالنسبة للنموذج الرياضي أو الجبري سواء كانت ن_١ = ن_٢ = = ن_ر أو أنها غير متساوية والمثال التالي يوضح ذلك .

مثال (١-٥) :

لاختبار الفرق بين أثر كل من خمس برامج تدريبية على الآلة الكاتبة ، فقد جربت البرامج الخمس على ٢٦ متدربا ، وبعد انتهاء فترة التدريب سجل الوقت الذي استغرق في كتابة تقرير معين لكل من المتدربين الـ ٢٦ (الوقت مقربا إلى أقرب دقيقة) وكانت النتائج كالآتي :

م	الوقت المستغرق في كتابة التقرير بالدقيقة (س)	
٣١٦	٨٥ ٨١ ٧٢ ٧٨	البرنامج أ
٤٤٣	٧٩ ٧١ ٨١ ٧٥ ٧٤ ٦٣	البرنامج ب
٥٠٥	٩٥ ٧٤ ٨١ ٨٦ ٩٠ ٧٩	البرنامج جـ
٥٨٦	٩٠ ٨٣ ٨١ ٧٩ ٧٥ ٩١ ٨٧	البرنامج د
٢٣٤	٨١ ٧٧ ٧٦	البرنامج هـ
٢٠٨٤=م		

مجموع $\sum R = 168368$ ، $\sum M = 167040.61$ ، $\sum N = 167040.61$

وتكون نتائج التحليل كالاتى :

مجموع المربعات الكلى = $168368 - 167040.61 = 1327.39$

مجموع المربعات الذى يعزى إلى اختلاف البرامج (أى مجموع المربعات بين

$$\text{البرامج} = \frac{(\sum R)^2}{N} + \frac{(\sum M)^2}{M} + \frac{(\sum N)^2}{N} = \frac{(316)^2}{4} + \frac{(443)^2}{6} + \frac{(505)^2}{7} + \frac{(586)^2}{7}$$

$$+ \frac{(234)^2}{3} - 167040.61 = 444.29$$

أما جدول تحليل التباين فيكون :

جدول تحليل التباين

المصدر	مجموع المربعات م.م	درجات الحرية د.ح	متوسط المربعات م.م	نسبة التباين ف.ف	ف.ف %
بين البرامج	444.29	4	111.07		
البواقي (أو الخطأ العشوائى)	883.10	21	42.05	2.64	2.84
الكلى	1327.39	25	-		

الفرض العدمى : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$

والفرض البديل هو أن الأثر المتوسط لأي برنامجين على الأقل غير متساو .
وهنا يقبل الفرض العدمي ، حيث $F = 2,64 > F_{0,01,4} = 2,84$

ملحوظة :

كان يمكن اختصار العمليات الحسابية بشكل واضح في هذا المثال أيضا بطرح وسط فرضي (٨٠ مثلا) من جميع القراءات وتحليل النتائج باستخدام القراءات المختصرة وسوف نحصل على نفس النتائج كما في جدول تحليل التباين الأخير.

(٣) العلاقة بين التصميم كامل العشوائية ، حيث $(J = 3)$ واختبار الفرض

$$\mu_1 = \mu_2$$

إذا كانت التجربة العشوائية عبارة عن تجربة أثر عامل واحد يتكون من عاملين $(J=2)$ فقط ، أي تجربة أثر معالجتين فقط ، فإن اختبار الفرض العدمي :

$\mu_1 = \mu_2$ مقابل الفرض البديل $\mu_1 \neq \mu_2$ ، يقوم على أساس مقارنة نسبة التباين من جدول تحليل التباين بالقيمة المناظرة من جدول توزيع ف بدرجات حرية ١ ، $N_1 + N_2 - 2 = N - 2$ عند مستوى المعنوية α ، ويرفض الفرض العدمي أو يقبل في ضوء نتيجة المقارنة .

وفي الحقيقة إن هذا الاختبار لا يختلف في النتيجة عن اختبار الفرق بين المتوسطين المعروف باختبار ت حيث :

الفرض العدمي : $\mu_1 = \mu_2$ أو $\mu_1 - \mu_2 = 0$ صفر

مقابل الفرض البديل $\mu_1 \neq \mu_2$

حيث يرفض الفرض العدمي متى كانت :

$$T = \frac{(\bar{S}_1 - \bar{S}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \frac{(S.D.)^2}{2 - N_1 + N_2}}} = *$$

$$= \frac{(\bar{S}_1 - \bar{S}_2) - (\bar{S}_3 - \bar{S}_4)}{\sqrt{\frac{1}{n} + \frac{1}{n}}}$$

$$\geq \frac{2}{\alpha} \cdot 2 - 2 + 1 = 1$$

$$\leq \frac{2}{\alpha} \cdot 2 - 2 + 1 = 1$$

حيث E^2 أى التباين التجميعى هى نفسها متوسط مربعات البواقى (أو الخطأ العشوائى) فى جدول تكميل التباين .

أو باختصار $|t| \leq 2 - 2 + 1 = 1$ $\frac{2}{\alpha}$ (11/1)

مثال (٦-١) :

وتطبيقاً لذلك فإن اختبار عدم اختلاف إنتاجية الآتين (أ ، ب) مثلاً من بيانات المثال (٤-١) يمكن أن يتم بأحد أسلوبين :

أسلوب تحليل التباين:

$$\text{مجموع المربعات الكلى} = 47^2 + 53^2 + 49^2 + \dots + 52^2 - \frac{(525)^2}{10} = 27765 - 27562,5 = 202,5$$

مجموع المربعات بين نوعى الآلة (أ ، ب) الذى يعزى إلى اختلاف النوع

$$A \text{ عن } B = \frac{1}{5} (280^2 + 245^2) - 27562,5 = 122,5 - 27685 =$$

$$122,5 - 27685 =$$

$$\text{إذن مجموع المربعات المتبقى} = 202,5 - 122,5 = 80$$

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين	ف . ا . ا . %
بين الآلات	122,5	1	122,5	12,25	5,32
المتبقى	80,0	8	10		
الكلى	202,5	9			

الفرض العدمي : $\mu_1 = \mu_2 = \mu_3 = \text{صفر}$

ويرفض الفرض العدمي حيث $F = 12,25 < F_{0,01,2,8} = 5,32$

وبأسلوب اختبار الفرق بين متوسطين :

بفرض استقلال العينتين وأن المتغير له توزيع معناد وأن σ^2 واحدة في مجتمع الدراسة فإن :

الفرض العدمي : $\mu_1 = \mu_2 = \mu_3 = \text{صفر}$

الفرض البديل : $\mu_1 \neq \mu_2$

ووسيلة الاختبار هي :

$$|t| = \frac{|\bar{s}_1 - \bar{s}_2| - (\bar{s}_1 - \bar{s}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma^2}}$$

ويرفض الفرض العدمي متى كانت $|t| < 2,306$ حيث $t_{0,01,8} = 2,306$

ومن البيانات نجد أن :

$$\bar{s}_1 = 49, \bar{s}_2 = 56$$

$$E^2 = \frac{\text{مجموع (مجموع (س ر - س ر) : ن) : ن}}{2 - n_1 + n_2} = 2,1$$

$$r = 1, 2, \dots, n_1 \text{ أو } n_2$$

$$10 = \frac{50 + 30}{8}$$

= متوسط مربعات البواقي في جدول تحليل التباين الأخير .

$$\therefore |t| = \frac{49 - 56 - \text{صفر}}{\sqrt{\left(\frac{1}{5} + \frac{1}{5}\right) 10}} = 3,5$$

وبالتالي يرفض الفرض العدمي .

وفى الواقع فإن :

$$t^2(0) = t^2(3,0) = 12,25 = \text{نسبة التباين فى جدول تحليل التباين ، وبالمثل}$$

$$t^2(0,8) = t^2(2,306) = 0,318 = \text{ف } 0,801\%$$

وهذه النتيجة صحيحة لأى قيمة لـ ف بدرجة حرية واحدة فى البسط ، أى أن
ف_١، ر_١ = α ، ت_١، ر_١ = α ، وهى العلاقة بين قيمة ت وقيمة ف الجدولية .

(٤) المقارنات الفردية :

أما وقد رفض الفرض العدمى بأن م_١ = م_٢ = م_٣ = صفر فى
مثال (٤/١) فإنه قد يعنينا أن نجيب على التساؤل التالى :

أى المعالجات تختلف فى أثرها معنوياً ؟ وللإجابة على ذلك فإن المعادلة (٣/١)
تصلح لإنشاء العلاقة التالية :

$$|\bar{m}_i - \bar{m}_j| \leq t_{\alpha, 2n-2} \times \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \times \sqrt{\frac{1}{2} \text{ ع } t^2(0,8)} \quad (12/1)$$

حيث هـ ≠ ٠ ، ن_د = ن_ر أو غير متساويان

وتعبر العلاقة (١٢/١) عن ما يسمى بالحد الأدنى للفروق المعنوية
(Least Significant Difference) أو باختصار (LSD) وهى تعنى أن أى
فرق بين أى متوسطين يساوى أو يتجاوز فى قيمته العددية الطرف الأيسر من
تلك العلاقة ، فإنه يعتبر فرقاً معنوياً.

وتطبيقاً لذلك على المعالجات التى اختبرت فى المثال (٤-١) نجد أن :

مثال (٧-١) :

$$t^2(0,8) = 0,318 \times \sqrt{\frac{1}{5} + \frac{1}{5}} \times \sqrt{\frac{1}{2} \text{ ع } t^2(2,306)} = 0,83 \sqrt{\frac{1}{5}} = 0,4$$

$$|\bar{m}_1 - \bar{m}_2| = 7, |\bar{m}_1 - \bar{m}_3| = 5, |\bar{m}_2 - \bar{m}_3| = 2$$

والفرقين الأول والثاني معنويين والأخير غير معنوي مما يعنى أن وجود ب فى أى من العلاقات الثلاث أدى إلى ظهور الفرق معنوياً ، ومثل هذه النتيجة تفيد فى اتخاذ القرارات.

وجدير بالذكر أننا ما كنا نستطرد فى التحليل للوصول إلى هذه المرحلة لو أننا قبلنا الفرض العدمى بتساوى أثر المعالجات.

وطريقة الحد الأدنى للفرق المعنوية تبدو وكأنها طريقة سهلة وسريعة للكشف عن الفروق المعنوية بين أى متوسطين ، ولكن يعاب عليها أنها تؤدي إلى زيادة مساحة منطقة الرفض (α) المقررة سلفاً بازدياد عدد المقارنات وهناك طرق أخرى تفضل طريقة الحد الأدنى للفرق المعنوية ويمكن للقارئ أن يستشير فى ذلك أحد المراجع المتخصصة منها ما ذكر بقائمة المراجع فى نهاية الكتاب .

(5) تعليق ختامى

(١-٥) يتميز النموذج كامل العشوائية بسهولة التحليل وبتوفير أكبر عدد ممكن من درجات الحرية للخطأ ، كما أن فقد أى مفردة أو مشاهدة - لسبب لا يتصل بالتجربة - لا يشكل صعوبة فى التحليل إذ يمكن إسقاطها من الحساب ولا حاجة إلى تقديرها ولا يؤثر ذلك كثيراً على دقة التقديرات ، وأن كان يعاب عليه عدم إمكان التحكم فى الخطأ العشوائى الذى يشمل جميع أنواع التباين عدا ما يرجع إلى المعالجات وهو ما تحاول معالجته النماذج الأخرى مثل نموذج القطاعات العشوائية.

(٢-٥) هذا وإذا لم يتحقق شرط من شروط صلاحية استخدام نموذج تحليل التباين فى حالة النموذج العشوائى الكامل ، فإنه يمكن استخدام أحد أساليب الإحصاء اللامعلمى كاختبار كروسكال - والس الذى سنتعرض له مستقبلاً.

تمارين

١- يعتقد أحد المحللين الماليين أن السهم من النوع أ يحقق عادة سعراً في التداول في سوق الأوراق المالية أعلا منه عن السهم من النوع ب . ومع ذلك فإن السهم من النوع أ يمثل مخاطرة أعلى في التقلبات اليومية للأسعار عنه بالنسبة للسهم من النوع ب وذلك على أساس التباين في تقلب الأسعار اليومية . وللتحقق من هذا فقد سجلت البيانات التالية عن عينة عشوائية من النوعين من الأسهم وذلك خلال يوم ما:

السهم من النوع		
ب	أ	
٢٥	٢٥	ن = ١
٠,١٢٥	٠,٢٥٠	متوسط التقلب في السعر س
٠,٤٦	٠,٧٦	ع

قارن بين درجة تقلب السعر اليومي لنوعى الأسهم أ ، ب وذلك باختبار الفرض العدمي بتساوى درجة تباين تقلب السعر اليومي عند مستوى المعنوية $\alpha = 5\%$.

٢- تعمل شركتان فى تجميع أجهزة التلفزيون . وقد تبين أنه خلال الأيام العشرة الأخيرة من شهر ديسمبر ٢٠٠٢ فقد أرتجع للشركة الأولى عدد ٩ أجهزة فى المتوسط بانحراف معيارى عدد ٢ جهاز . فى حين أنه بالنسبة للشركة الثانية وخلال نفس الفترة كان متوسط عدد الأجهزة المرتجعة عدد ٨,٥ جهاز بانحراف معيارى قدره ١,٥ جهاز . هل تشير هذه النتائج إلى أن تباين توزيع المرتجع من الأجهزة من انتاج الشركة الأولى أعلى منه للشركة الثانية عند $\alpha = 5\%$ ؟

٣- أ : سجلت قراءات عشوائية عددها $n = 100$ باستخدام آلة قياس وكانت

نتيجة هذه التجربة كالتالي: $\bar{x} = 9.4$ وحدة قياس ، $s^2 = 4.84$

اختبر الفرض العدمي بأن $\sigma^2 = 1$ مقابل الفرض البديل $\sigma^2 < 1$ عند

مستوى المعنوية $\alpha = 0.05$

ب : ماذا لو كان عدد القراءات $n = 7$ ، $\bar{x} = 4.74$ ؟

فماذا سيكون القرار في هذه الحالة ؟

٤- الجدول التالي يعرض بيانات جزئية لتحليل التباين:

المصدر	م.م	د.ح	م.م	ف
بين الأنواع أو العينات		٢		
داخل العينات أو البواقي			٢٠	
كلى	٥٠٠	١١		

أ : أكمل جدول تحليل التباين . كم عدد الأنواع موضوع المقارنة بهذه

التجربة ، وما هو حجم العينة الكلى.

ب: حدد الفرض موضوع الاختبار (الفرض العدمي) والفرض البديل ثم

اختبر الفرض عند $\alpha = 5\%$.

٥- سحبت عينتين عشوائيتين مستقلتان كل من مجتمع له التوزيع المعتاد لهما

المتوسط والتباين (μ_1, σ_1^2) ، (μ_2, σ_2^2) على التوالي . وكانت النتائج

التي سجلت على العينتين كالتالي :

العينة الأولى	العينة الثانية
$n_1 = 20$	$n_2 = 15$
$\bar{x}_1 = 123$	$\bar{x}_2 = 116$
$s_1^2 = 31.3$	$s_2^2 = 120.1$

أ: اختبر الفرض العدمي بأن $\sigma^2 = \sigma^2$ مقابل الفرض بأن $\sigma^2 \neq \sigma^2$
استخدم $\alpha = 0.05$

ب: هل يمكنك استخدام اختبارات لاختبار الفرض العدمي بأن $(\mu_1 - \mu_2) = 0$ مقابل الفرض بأن $(\mu_1 - \mu_2) \neq 0$ ؟ لماذا ؟

٦- الجدول التالي يبين عدد الوحدات المنتجة في خمسة أيام متتالية بواسطة أفراد عينة عشوائية من العمال عددهم ٢٠ عاملاً .

	الأيام				
	١	٢	٣	٤	٥
	٣٠	٣٠	٤٠	٤٥	٥٥
	٣٥	٤٠	٤٥	٤٠	٤٥
	٢٥	٤٥	٥٥	٣٥	٦٠
	٤٠	٤٥	٣٥	٤٠	٥٠
المجموع	١٣٠	١٦٠	١٧٥	١٦٠	٢١٠

استخدم أسلوب تحليل التباين في تحليل نتائج هذه التجربة عند $\alpha = 5\%$.
أوجد الخطأ المعياري للفرق بين أي متوسطين واستخدم ذلك في إنشاء فترة
ثقة للفرق بين أي متوسطين عند $\alpha = 5\%$ وفي ضوء ذلك ناقش الفروق
بين متوسطات الإنتاج في الأيام الخمسة .

٧- البيانات التالية تبين عدد الشيكات التي صرفت خلال كل يوم عمل في عينة
من أربعة فروع لأحد المصارف :

	٢٤٥	٢٦٣	٢٦٦	٢٤٩	٣٣٧	٢٧٧	٢٨٩	٢٤٤	٢٦٥
أ									
ب	٢٣٦	١٩١	٢٠٩	٢٥٢	٢١٢	٢٢٤			
ج	٢٥٣	١٩٢	١٤١	١٦٠	٢٢٩	٢٢١	١٥٠	٢١٥	٢٣٤
د	١٧٣	١٦٤	١٨٣	١٣٨	١٤٦	١٢٥	١٧٨	١٩٩	١٧٠
	١٨٨	١٩٨							

حلل النتائج السابقة باتباع أسلوب تحليل التباين عند مستوى المعنوية ٥% .
أوجد مجموع المربعات للخطأ العشوائي ثم بطريقة الجمع من العينات
احسب قيمة معامل الاختلاف النسبي حيث معامل الاختلاف النسبي =

$$100 \times \frac{C}{S}$$

٨- فيما يلي عدد الساعات التي استغرقت في إنتاج أربعة أنواع من سلعة ما
(بعد طرح ١٠٠ ساعة) .

السلعة	١	٢	٣	٤	
	٦٤	٧٨	٧٥	٥٥	
	٧٢	٩١	٩٣	٦٦	
	٦٨	٩٧	٧٨	٤٩	
	٧٧	٨٢	٧١	٦٤	
	٥٦	٨٥	٦٣	٧٠	
	٩٥	٧٧	٧٦	٦٨	
م.م	٤٣٢	٥١٠	٤٥٦	٣٧٢	١٧٧٠ = م
م.س.ر	٣١٩٩٤	٤٣٦٥٢	٣٥١٤٤	٢٣٤٠٢	

حلل هذه البيانات مستخدماً أسلوب تحليل التباين عند مستوى معنوية ٥% .
٩- البيانات التالية تبين عدد من تغيب عن العمل دون إذن مسبق خلال نصف
سنة في ثلاثة وحدات إنتاجية :

	الوحدة الإنتاجية			عدد أيام التغيب دون إذن مسبق
	ج	ب	أ	
	٣	١	٦	٢
	٥	٣	٤	٣
	٥	٣	٩	٤
	٨	٦	٨	٥
	١٩	٦	١	٦
	٢٣	١٤		٧
	٢٢	١١		٨
	١٤	٤		٩
	١٤	٦		١٠
	٧	٢		١١
	٨	٣		١٢
	٤	١		١٣
	١			١٤
٢٢٤	١٣٣	٦٠	٣١	ر
١٦٠٤	١٠٣٧	٤٤٢	١٢٥	م.ن
١٣١٢٤	٨٩٦١	٣٦٠٢	٥٦١	م.س

وإذا علم أن $E^1 = ١,٩٠$ ، $E^2 = ٥,٨٦$ ، $E^3 = ٦,٦٤$ هي قيم التباين لتوزيع التغيب في الوحدات الإنتاجية الثلاثة . أوجد الخطأ المعياري للفرق بين متوسط طول مدة التغيب في الـ ١ ، ٢ ، ٣ باستخدام تقدير التباين التجميعي E^2 ثم باستخدام أسلوب تحليل التباين .

احسب نفس القيمة وعل اختلاف القيمتين إن وجد . ثم قدر الفرق بين متوسطى مدة التغيب فى الودعتين بفترة ثقة ٩٥% بالطريقتين .

١٠- فى دراسة عن درجة اختلاف الأسر ذات الحجم الواحد من سكان المناطق السكنية (حضر - ريفية محضرة - ريفية) فى الإنفاق على المأكّل والمشرب خلال فترة الدراسة مخفضا بـ ٢٥٠ ج فكانت النتائج كالاتى :

ح = س - ٢٥٠

١١٤	٩٦	١٠١	١٥١	٨١	مناطق حضرية
	١٩	٢٢	٢٥	١٢	مناطق ريفية محضرة
		٩-	١٨	٤١-	مناطق ريفية

هل تؤيد هذه النتائج الاعتقاد بعدم اختلاف الأسر فى إنفاقها على المأكّل والمشرب باختلاف محل إقامتها عند مستوى المعنوية ٥% ؟

١١- استخدم أسلوب تحليل التباين فى اختبار أثر مدة الخبرة على إنتاجية العامل يوم / جنيه من البيانات التالية :

س : الإنتاجية يوم / جنيه					عدد سنوات الخبرة	عدد العمال فى كل مجموعة
٢٧	٣١	٣٢	٣٧	٣٣	٥	٨
		٢٨	٣٥	٣٠	٣	١٢
	٣٦	٣٤	٢٩	٣٣	٤	١٦

استخدم وسطا فرضيا و = ٣٠ فى تبسيط عملياتك الحسابية .

الفصل الثاني

التصنيف متعدد الاتجاهات

محتويات الفصل

مقدمه

أولاً: تصميم القطاعات الكاملة العشوائية:

(١) القطاعات الكاملة العشوائية بدون تكرار.

(١-١) النموذج الرياضى والنموذج الحسابى.

(٢-١) طبيعة حد البواقي.

(٣-١) الكفاءة النسبية للنموذج.

(٤-١) القراءات المفقودة.

(٥-١) الخطأ المعيارى للفرق بين متوسطين.

(٢) القطاعات الكاملة العشوائية مع التكرار.

(١-٢) مقدمه

(٢-٢) النموذج الرياضى.

(٣-٢) نموذج جدول تحليل التباين لتصميم القطاعات الكاملة العشوائية

مع التكرار (النموذج الثابت).

ثانياً : المربع اللاتيني:

(١) مقدمه.

(٢) النموذج الرياضى والحسابى.

(٣) الكفاءة النسبية لنموذج المربع اللاتيني.

(٤) القراءات المفقودة.

ثالثاً : التحليل العاملى:

(١) مقدمة : الأثر الأساسى والتفاعل.

(٢) التحليل العاملى لتجربة (٢×٢).

(٢-١) النموذج الرياضى.

تمارين

الفصل الثاني

التصنيف متعدد الاتجاهات

Multiway Classification

مقدمة:

كان المصدر الوحيد للتباين في قيم وحدات التجربة هي المعالجات، ولذلك فقد تعرضنا في الباب السابق إلى تصنيف المعالجات بحسب أنواعها أو مستوياتها وبالتالي كانت الخاصية التي تحكم التصميم تفرض اتجاهًا واحدًا لتوزيع المعالجات سواء كان هذا الاتجاه أفقيًا (أي في شكل صفوف) أو رأسيًا (أي أعمدة). ثم يتم توزيع وحدات التجربة عشوائيًا على جميع أنواع أو مستويات المعالجات، أي أن أسلوب تجميع وحدات التجربة في مجموعات - كل تمثل نوعاً أو مستوى للمعالجة - هو التجمع أحادي العنصر (Single Grouping).

وننتقل الآن إلى توزيع المعالجات بحسب خاصية أخرى، سنطلق عليها اسم القطاعات (Blocks). ونقدم في هذا الصدد نموذج القطاعات الكاملة العشوائية Randomized Complete Blocks Design. ثم ننتقل بعد ذلك إلى توزيع المعالجات بحسب خاصيتين تخصص لأحدهما الصفوف والأعمدة للخاصية الثانية وهو التصميم المعروف باسم (Latin Square) وكلاهما يدخل في إطار النماذج متعددة التصنيفات أو الاتجاهات (Multiway Classification). ثم نختم معالجتنا لهذا الموضوع بتقديم محدود لنموذج التحليل العاملي، نقتصر فيه على نموذج التحليل العاملي 2×2 .

أولاً : تصميم القطاعات الكاملة العشوائية Randomized Complete Blocks Design

(١) القطاعات الكاملة العشوائية بدون تكرار:

لنفرض أن لدينا (ر ل) وحدة تجريبية قسمت إلى ر قطاع (Block) بكل ل معالجة (Treatment) وإن كل معالجة خصصت عشوائياً لكل من الوحدات التجريبية في كل قطاع ، فإن تجربة كهذه تعرف باسم القطاعات الكاملة العشوائية. وفي مثل هذا النموذج يكون كل قطاع متجانساً قدر الإمكان، ولكن تتباين القطاعات فيما بينها. ويكون عدد الوحدات التجريبية التي تخصص لكل معالجة في كل قطاع متساوية في العدد (عادة وحدة واحدة في كل معالجة في كل قطاع) ومتى تحقق ذلك فإن الفروق بين القطاعات تكون متعامدة مع أثر المعالجات. ويتبع هذا الأسلوب عادة لاستبعاد أثر عامل ثان هو القطاعات ، هذا العامل قد يكون له أثره بالإضافة إلى العامل الأول وهو المعالجات.

ففى مثال (١-٢) فى الفصل السابق ، يمكن أن تكون الأعمدة تمثل العمل فى خمسة أيام متتالية أو فى خمسة وحدات مختلفة ، تعبر عن القطاعات فى حين تعبر الصفوف عن أنواع الآلات الثلاثة أى أن ناتج التجربة يكون كالتالى:

م ل	القطاعات (الوحدات الإنتاجية)					المعالجات (الآلة)
	أ	ب	ج	د	هـ	
٢٤٥	٤٧	٥٣	٤٩	٥٠	٤٦	(١)
٢٨٠	٥٥	٥٤	٥٨	٦١	٥٢	(٢)
٢٥٥	٥٤	٥٠	٥١	٥١	٤٩	(٣)
٧٨٠	١٥٦	١٥٧	١٥٨	١٦٢	١٤٧	م ر

وكما يتبين من هذا الجدول ، فإن كل معالجة تظهر في جميع القطاعات وبنفس العدد كما أن كل قطاع يشتمل على جميع المعالجات بنفس عدد الوحدات التجريبية (وحدة واحدة في كل قطاع في كل معالجة في مثالنا هذا).

(1-1) النموذج الرياضي أو النموذج الجبري (أو الحسابي):

ويمثل هذا النموذج بالمعادلة التالية:

$$Y_{rj} = \mu + \alpha_r + \beta_j + \gamma_{rj} + \epsilon_{rj} \quad (1/2)$$

ويفترض في هذا النموذج تخصيص وحدة تجريبية واحدة عشوائيا لكل قطاع وكل معالجة، أي دون تكرار (Single replication) والفروض الخاصة بالنموذج (1/2) لا تختلف في شيء عن فروض النموذج (1/1) إلا في إضافة الحد γ_{rj} للدلالة على الآثار المضافة للقطاع r ($r = 1, 2, \dots, L$).

ويتكون مجموع المربعات الكلي من مركبات ثلاث ، المركبة الأولى تعبر عن أثر المعالجات والمركبة الثانية تعبر عن أثر القطاعات أما الثالثة فتشير إلى الخطاء العشوائي (أو البواقي) أي التباين الذي لا يفسر على أساس من المعالجات أو القطاعات ، ولكن يرجع إلى عامل الصدفة أو عوامل لا تتكرر بانتظام ويصعب التحكم فيها. أي أن:

مجموع المربعات الكلي = مجموع المربعات بين المعالجات (أو الذي يعزى إلى اختلاف أثر المعالجات) + مجموع المربعات بين القطاعات (أو الذي يعزى إلى اختلاف أثر القطاعات) + مجموع المربعات المتبقية (ويعزى إلى الخطاء العشوائي). (2/2)

أى أن:

$$\begin{aligned} & \text{مجموع } (س.د. - س.د.)^2 = \text{مجموع } (س.د. - س.د.)^2 \\ & + \text{مجموع } (س.د. - س.د.)^2 + \text{مجموع } (س.د. - س.د.)^2 \\ & - (س.د. + س.د.)^2 \end{aligned} \quad (3/2)$$

وعادة ما تستخدم الصيغة التالية لحساب مجموع المربعات:

$$\begin{aligned} & \text{مجموع } (س.د. - س.د.)^2 = \frac{\text{مجموع } (س.د.)^2}{س.د.} - \frac{\text{مجموع } (س.د.)^2}{س.د.} \\ & + \left(\frac{\text{مجموع } (س.د.)^2}{س.د.} - \frac{\text{مجموع } (س.د.)^2}{س.د.} \right) \end{aligned} \quad (4/2)$$

حيث م : مجموع قيم الوحدات التجريبية = مجموع س.د. من س.د.

$$\text{س.د. المتوسط العام} = \frac{1}{س.د.} = \frac{م}{س.د.} = \frac{م}{س.د.}$$

ر : مجموع قيم الوحدات التجريبية للمعالجة ل.

س.د. : متوسط قيم الوحدات التجريبية للمعالجة ل.

ر : مجموع قيم الوحدات التجريبية للقطاع ر.

س.د. : متوسط قيم الوحدات التجريبية للقطاع ر.

ويستكمل تحليل النتائج باستخدام جدول تحليل التباين كالاتي:

المصدر	مجموع المربعات (م)	درجات الحرية (د.ح)	متوسط المربعات م.م.م	نسبة التباين ف
بين المعالجات	$\frac{1}{r} \sum_{j=1}^r r_j^2 - \frac{(\sum_{j=1}^r r_j)^2}{l \times r}$	$l - 1$	$\frac{م.م.م}{l - 1}$ (1) =	(1) ÷ (3)
بين القطاعات	$\frac{1}{l} \sum_{i=1}^l l_i^2 - \frac{(\sum_{i=1}^l l_i)^2}{r \times l}$	$r - 1$	$\frac{م.م.م}{r - 1}$ (2) =	(2) ÷ (3)
البواقي	بالفرق	$(l-1)(r-1)$ (أو بالفرق) (3) =	$\frac{م.م.م}{(l-1)(r-1)}$ (3) =	-
الكل	$\sum_{j=1}^r r_j^2 - \frac{(\sum_{j=1}^r r_j)^2}{r \times l}$	$rl - 1$	-	-

أولاً : الفرض العدمي H_0 : $\mu_1 = \mu_2 = \dots = \mu_r = \mu$ = صفر

والفرض البديل هو اختلاف الأثر المتوسط لأي معالجتين على الأقل ، ويرفض

الفرض العدمي إذا كانت:

$$F = \frac{\frac{م.م.م}{l-1}}{\frac{م.م.م}{(l-1)(r-1)}} \leq F_{\alpha}(l-1)(r-1)$$

ويقبل الفرض العدمي في غير ذلك من الأحوال.

ثانياً: الفرض العدمي $Q_1 = Q_2 = \dots = Q_r = \text{صفر}$

والفرض البديل هو اختلاف الأثر المتوسط لأي قطاعين على الأقل ويرفض الفرض العدمي إذا كانت:

$$F = \frac{M \cdot Q}{1 - r} \div \frac{M \cdot X}{(1 - l)(1 - r)} \leq F_{\alpha, (1 - l)(1 - r), (1 - r)}$$

ويقبل الفرض العدمي في غير ذلك من الأحوال.

حيث ترمز م.م. إلى مجموع المربعات للمعالجات، م.ق إلى مجموع المربعات للقطاعات أما م.م.خ فتشير إلى مجموع المربعات للخطأ العشوائي أو البواقي .

وسوف نستخدم بيانات الجدول السابق في عرض طريقة تحليل النتائج لتجربة قطاعات عشوائية كاملة بدون تكرار.

مثال (١-٣)

الآلات	الوحدات الإنتاجية					م.ر	م.ج
	أ	ب	ج	د	هـ		
(١)	٤٧	٥٣	٤٩	٥٠	٤٦	٢٤٥	٤٩
(٢)	٥٥	٥٤	٥٨	٦١	٥٢	٢٨٠	٥٦
(٣)	٥٤	٥٠	٥١	٥١	٤٩	٢٥٥	٥١
م.ر	١٥٦	١٥٧	١٥٨	١٦٢	١٤٧	٧٨٠	٥٢
م.ج	٥٢	٥٢,٣٣	٥٢,٦٧	٥٤	٤٩	م	م.ج

وبفرض أن بيانات هذا الجدول تمثل إنتاج أفراد عينة عشوائية مكونة من ١٥ عاملاً ، وزعوا عشوائياً على وحدات إنتاجية عددها خمسة ، وبكل ثلاثة أنواع مختلفة من الآلات فإن تحليل التباين يظهر التالي:

$$\text{مجموع المربعات الكلية (م.م.ك)} = \sum_{\text{رد}} \sum_{\text{م.م.ك}} \frac{\sum_{\text{رد}}^2}{\text{رد} \times \text{رد}} - \frac{\sum_{\text{رد}}^2}{\text{رد}}$$

$$= 780^2 - 249 + \dots + 53 + 47 =$$

$$224 = 4060 - 40784 =$$

مجموع المربعات بين الآلات (الذي يعزى إلى اختلاف الآلات)

$$= \sum_{\text{رد}} \sum_{\text{م.م.ك}} \frac{\sum_{\text{رد}}^2}{\text{رد} \times \text{رد}} - \frac{\sum_{\text{رد}}^2}{\text{رد}}$$

$$= \left[(52 - 51) + (52 - 56) + (52 - 49) \right] 5 =$$

$$= 130 = 26 \times 5 =$$

$$\text{وهي أيضاً} = \frac{1}{\text{رد}} - \left[\sum_{\text{رد}} \frac{\sum_{\text{رد}}^2}{\text{رد} \times \text{رد}} - \frac{\sum_{\text{رد}}^2}{\text{رد}} \right]$$

$$= \frac{780^2}{15} - \left[(2500 + 2280 + 2450) \right] \frac{1}{5} =$$

$$= 4060 - \frac{6020 + 7840 + 6020}{5} =$$

$$= 4060 - 4069 = 130 =$$

مجموع المربعات بين الوحدات الإنتاجية (الذي يعزى إلى اختلاف الوحدات

$$\text{الإنتاجية) = } \sum_{\text{رد}} \sum_{\text{م.م.ك}} \frac{\sum_{\text{رد}}^2}{\text{رد} \times \text{رد}} - \frac{\sum_{\text{رد}}^2}{\text{رد}}$$

$$= \left[(52 - 52,67) + (52 - 52,33) + (52 - 52) \right] \times 3 =$$

$$= 13,0078 \times 3 = \left[(52 - 49) + (52 - 54) + \right]$$

$$= 40,6734 =$$

$$\text{وهي أيضاً} = \frac{1}{J} - \left[\frac{r^2}{(r^2)} \right] - \frac{r^2}{J \times J}$$

$$= \frac{1}{3} - \left[\left(\frac{1147}{15} + \frac{1162}{15} + \frac{108}{15} + \frac{107}{15} + \frac{106}{15} \right) \right]$$

$$= \frac{1}{3} - \left(\frac{2116.9 + 2624.4 + 2496.4 + 2464.9 + 2433.6}{3} \right) = 40.6734 = 40.67\%$$

ويكون مجموع المربعات المتبقى ويعزى إلى الخطأ العشوائى هو:

$$= \text{م.م.ك.} - \text{م.م.} - \text{بين الآلات} - \text{م.م. بين الوحدات الإنتاجية.}$$

أما جدول تحليل التباين فيكون كالاتى:

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بين الآلات	130,00	2	65	9,75
بين الوحدات الإنتاجية	40,67	4	10,17	1,52
البواقي	53,33	8	6,67	
الكل	224,00	14		

الفرض العدمى: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \text{صفر}$

والفرض البديل: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5$ لأى معالжин على الأقل.

ويرفض الفرض عند $\alpha = 5\%$ حيث $F_{0.05, 4, 8} = 4,46$ أى أنه لاختلاف

أنواع الآلات أثره المعنوى على وحدات التجربة عند مستوى المعنوية 5% .

أما الفرض العدمي : $Q_1 = Q_2 = Q_3 = Q_4 = Q_5 = 0$ - صفر

والفرض البديل: $q \neq r$ ، $h \neq$ و لأى قطاعين على الأقل .

وهنا يقبل الفرض البديل حيث $F = 3.84 > 1.54$ - ف $\alpha = 0.05$ ، أى أن

اختلاف الوحدات الإنتاجية ليس لها أثر على الوحدات التجريبية.

وفى حالة كهذه ، يمكن إعادة إدماج مجموع المربعات ودرجات الحرية

للوحدات الإنتاجية (طالما أنه تبين عدم معنوية أثرها) إلى البواقي، وبذلك نعود

إلى التحليل في اتجاه واحد ، حيث تتضح عدم جدوى اختلاف الوحدات الإنتاجية

في تفسير جزء من اختلاف الوحدات التجريبية في القيمة.

(١-٢) طبيعة حد البواقي:

سبق أن عَرَفَ مجموع المربعات للخطأ العشوائى أو مجموع المربعات للبواقي

بأنه:

(5/2) $\mu = (r_1 - \bar{r}_1, r_2 - \bar{r}_2, \dots, r_n - \bar{r}_n)$

حيث \bar{X}_1 ، \bar{X}_2 ، \bar{X}_3 هي متوسطات المعالجة ل و لقطاع ر والمتوسط

العام على التوالي ، ومجموع المربعات طبقاً لهذا التعريف الذي يقرر باستخدام

طريقة المربعات الصغرى ، يؤدي إلى الحصول على أقل مجموع لمربعات

الخطأ الذي يحقق أحسن تقدير لـ μ ، σ ، ρ وهي معالم النموذج:

(۱/۲) $\mu_{-} + \mu_{-} + \mu_{-} + \mu_{-}$

حيث يعبر عن هذا النموذج في صورة بيانات العينة على النحو التالي:

$$\bar{s}_r = \bar{s}_{r-1} + (\bar{s}_{r-1} - \bar{s}_{r-2}) + (\bar{s}_{r-2} - \bar{s}_{r-3}) + \dots + \bar{s}_1$$

$$(6/2) \quad \mu + \alpha + \beta + \gamma =$$

حيث يعرف التأثير المضاف للمعالجات $\mu^A = (\bar{S}_N - \bar{S}_{..})$ (٧/٢)

والتأثير المضاف للقطاعات $Q = (S_r - S_{..})$ (٨/٢)

بالفرق بين متوسط المعالجة (ل) والوسط العام ، وكذلك متوسط القطاع (ر) والمتوسط العام. وحيث:

$$\bar{e}_r = (\bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..}) = \text{صفر}$$

ومن (١/٢ ، ١) يتبين الآتي:

$$\bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..}$$

$$(١/٢) \quad \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..}$$

وباستخدام أسلوب المربعات الصغرى يمكن تقدير قيم $\mu = \bar{e}_{..}$ ،

$$\bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..}$$

وهي تقديرات معالم النموذج (١/٢) بقيم وحيدة وهذه التقديرات هي:

$$(١٠/٢) \quad \left\{ \begin{array}{l} \bar{e}_{..} = \frac{\mu}{r \times l} \\ \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} \\ \bar{e}_{..} = \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} - \bar{e}_{..} \end{array} \right.$$

(٣-١) الكفاءة النسبية للنموذج:

كثيراً ما يعني أن نتعرف على مدى كفاءة استخدام نموذج القطاعات الكاملة العشوائية ، كبديل للنموذج العشوائي الكامل ، أي ما هو العائد من تخصيص ل وحدة تجريبية عشوائية في كل قطاع ، بدلا من توزيع ر ل وحدة عشوائية على ل معالجة ، ويتمثل ذلك في تخفيض خطأ التجربة أو الخطأ العشوائي بطريقة واضحة. ولتحقيق ذلك يستخدم معامل الكفاءة النسبية (Efficiency) حيث:

م.ك. ن (قطاعات عشوائية / عشوائي كامل) = $\frac{\sum_{i=1}^n \sum_{j=1}^m x_{ij}^2}{\sum_{i=1}^n \sum_{j=1}^m x_{ij}}$ (١١/٢)
 حيث ع_٢: متوسط مربعات الخطأ العشوائي للنموذج كامل العشوائية.
 ع_٢: متوسط مربعات الخطأ العشوائي لنموذج القطاعات العشوائية.
 ونظراً لأنه لا يتوفر لدينا في هذه الحالة إلا ناتج التجربة لنموذج القطاعات العشوائية الكاملة ، لذلك تعين علينا تقدير ع_٢ من جدول تحليل التباين لهذا النموذج الأخير وهذا التقدير هو:

$$\text{ع}_2 = \frac{\sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2}{n-1} \quad (١٢/٢)$$

وبالتالي فإن :

$$\text{م.ك. ن.} = \frac{\sum_{i=1}^n \sum_{j=1}^m x_{ij}^2}{\sum_{i=1}^n \sum_{j=1}^m x_{ij}} - \frac{\sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2}{\sum_{i=1}^n \sum_{j=1}^m x_{ij}} \quad (١٣/٢)$$

وجدير بالذكر أن تقدير ع_٢ من نتائج تحليل بيانات قطاعات عشوائية كاملة ≠ (مجموع مربعات القطاعات + مجموع مربعات البواقي) ÷ ل (ر-١) أي بإدماج مجموع مربعات القطاعات ودرجات حرية القطاعات إلى مجموع مربعات ودرجات حرية البواقي ، بفرض أن النموذج من البداية هو نموذج قطاعات عشوائية ، لأن ذلك إن صح ، فإن مجموع المربعات للمعالجات سوف يكون واحدا سواء كان النموذج المتبع في تصميم التجربة ، هو نموذج العشوائية الكاملة الذي تخصص فيه ر ل وحدة تجريبية على جميع المعالجات عشوائياً أو نموذج القطاعات الكاملة العشوائية حيث توزع ل وحدة عشوائية داخل كل قطاع، وذلك لأن مجموع المربعات الكلي واحد في الحالتين واختلاف التصميم لا بد وأن يؤدي إلى اختلاف المجموع للمعالجات.

وعادة ما يصحح معامل الكفاءة النسبي (١٣/٢) وذلك للتصحيح مقابل نقص درجات الحرية للخطأ العشوائي أو البواقي باستخدام نموذج القطاعات الكاملة العشوائية وذلك على النحو التالي:

$$م.ك.ن = \frac{٢٤}{٢٤} \times \frac{(١ + ٢ن) (٣ + ١ن)}{(١ + ١ن) (٣ + ٢ن)} \times ١٠٠ \times \frac{(١٤/٢)}{٢٤}$$

حيث

ن١: درجات حرية البواقي — ع٢ أى باستخدام النموذج العشوائي الكامل.
ن٢: درجات حرية البواقي — ع٢ أى باستخدام نموذج القطاعات الكاملة العشوائية.

ولقياس كفاءة استخدام القطاعات الكاملة العشوائية في المثال الأخير فإن:

معامل الكفاءة النسبي (ع / ق ع)

$$١٠٠ \times \frac{(٣ + ١٢) (١ + ٨)}{(١ + ١٢) (٣ + ٨)} \times \frac{٤٠,٦٧ + ٦,٦٧ \times ٢ \times ٥}{٦,٦٧ \times ١٤} =$$

$$١٠٠ \times \frac{١٥ \times ٩}{١٣ \times ١١} \times \frac{٤٠,٦٧ + ٦٦,٧}{٩٣,٣٨} =$$

$$١٠٠ \times ٠,٩٤٤ \times ١,١٤٩٨ = ١٠٨,٥ \%$$

وهذا يعنى أن الخطأ العشوائي لو استخدم تصميم العشوائية الكاملة ووزعت الـ ١٥ وحدة تجريبية عشوائياً على ثلاث معالجات ، لكان ١٠٨% من قيمته لو استخدم أسلوب القطاعات العشوائية بتوزيع ثلاث وحدات تجريبية عشوائياً على كل قطاع من القطاعات الخمسة ، وهذا يؤدي بالطبع إلى زيادة فترة الحد الأدنى للفروق المعنوية اللازمة لاختبار معنوية الفرق بين أى متوسطين ، كما أن ذلك يعنى أن استخدام ١٠٨ وحدة تجريبية باستخدام النموذج كامل العشوائية ، سوف يؤدي إلى الوصول إلى نفس الدرجة من الدقة باستخدام ١٠٠ وحدة فقط مع التحليل في اتجاهين باتباع نموذج القطاعات العشوائية ،

وتزداد كفاءة القطاعات العشوائية كلما ازدادت معنوية أثر القطاعات كما سيبتين من المثال التالي.

مثال (٣-٣) :

البيانات التالية تبين نتائج تجربة استخدم فيها نموذج القطاعات الكاملة العشوائية ، حيث اعتبرت أنواع الدهون (٨) معالجات ، اختبرت جميعها في كل يوم من الـ (٦) أيام وهي تمثل القطاعات ، وكانت التجربة لقياس تباين أو اختلاف نوع من الفطائر في درجة امتصاصه للدهون بحسب النوع واليوم كما يلي:

م.ج	(نوع الدهون) معالجات								الأيام (قطاعات)
	٨	٧	٦	٥	٤	٣	٢	١	
١٣٣١	١٦٤	١٥٠	١٦٣	١٦٣	١٧٨	١٧٧	١٧٢	١٦٤	(١)
١٤٧٢	١٦٩	١٧٩	١٩٣	١٧٧	١٩٦	١٨٤	١٩٧	١٧٧	(٢)
١٣٢٠	١٥٥	١٤٦	١٧٦	١٤٤	١٧٧	١٨٧	١٦٧	١٦٨	(٣)
١٢٩٤	١٤٩	١٤١	١٧٢	١٦٥	١٨١	١٦٩	١٦١	١٥٦	(٤)
١٣٩٦	١٧٠	١٦٩	١٧٦	١٦٦	١٨٤	١٧٩	١٨٠	١٧٢	(٥)
١٤٧٩	١٦٧	١٨٣	١٧٨	١٧٨	١٩١	١٩٧	١٩٠	١٩٥	(٦)
م - ٨٢٩٢	٩٧٤	٩٦٨	١٠٥٨	٩٩٠	١١٠٧	١٠٩٣	١٠٦٧	١٠٣٢	م.ج

$$م.م.ك = \frac{٨٢٩٢^2}{٤٨} - (١٦٧^2 + \dots + ١٧٢^2 + ١٦٤^2)$$

$$= ٩١٤٣,٠٠ - ١٤٣٢٤٤٣ - ١٤٤١٥٨٦ =$$

* Anderson, R.L. and Bancroft T.A. Statistical Theory In Research, Mcgraw-Hill Book Company, 1952, pp. 238-239.

م.م. بين الدهون (معالجات) =

$$= \frac{1}{7} (1032 + 1067 + \dots + 974) - 1432443 = 3344,33$$

م.م. بين الأيام (القطاعات) =

$$= \frac{1}{8} (1331 + 1472 + \dots + 1479) - 1432443 = 3986,75$$

$$\text{م.م. المتبقى} = 9143 - 3344,33 - 3986,75 = 1811,92$$

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
بين الأيام (ق)	3986,75	5	797,35	15,40
بين الدهون (م)	3344,33	7	477,76	9,23
الباقى	1811,92	35	51,77	—
كلى	9143,00	47	—	—

الفرض العدمي: $\mu_1 = \mu_2 = \dots = \mu_8 = \text{صفر}$

الفرض العدمي: $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_7^2 = \text{صفر}$

والفرض البديل في كل حالة: هو عدم تساوى الأكثر المتوسط لأى نوعين على الأقل من الدهون (المعالجات) ، وكذلك للأثر المتوسط ليومين على الأقل (القطاعات).

ولكن ف $1.35.0\%$ ، $3.70 >$ ، ف $1.35.7\%$ ، $3.3 >$ لذلك يرفض
الفرضين عند مستوى معنوية 1% .
في هذا المثال سوف نجد أن:
معامل الكفاءة النسبي (ق ع / ع)

$$= \frac{3986.75 + 51.77 \times 24}{51.77 \times 47} \times \frac{43 \times 36}{41 \times 38} - 100 \times \frac{43 \times 36}{41 \times 38} = 251.5\%$$

مما يعني أن استخدام نموذج العشوائية الكاملة بتوزيع الدهون عشوائيا
على الأيام الستة لكان الخطأ العشوائي (متوسط مربعات البواقي) قد ارتفع إلى
 251% من قيمته الحالية أو أننا كنا سنستخدم عددا من الوحدات التجريبية يصل
إلى مرتين ونصف عدد الوحدات التي استخدمت في هذه التجربة للوصول إلى
نفس القيمة للخطأ العشوائي.

(1-2) القراءات المفقودة: Missing Plots:

قد تفقد إحدى القراءات أو أكثر نتيجة عوامل خارجة عن الإرادة ولا
تتصل بظروف التجربة. وإذا كان النموذج المستخدم هو النموذج العشوائي
الكامل ، فإن فقد قراءة أو أكثر لا يؤثر على التحليل من حيث الأسلوب متى كان
عدد القراءات المتبقية لكل معالجة لا يقل عن 2 ، إلا أن درجات الحرية للخطأ
تنقص درجة مقابل كل وحدة تجريبية تفقد ، ويتم التحليل كالمعتاد سواء كان عدد
الوحدات لكل معالجة متساو أو غير متساو . وجدير بالذكر أن تعدد القراءات
المفقودة يستدعي مراجعة تخطيط التجربة وكيفية تنفيذها.

أما إذا كنا بصدد تجربة قطاعات كاملة عشوائية ، فإن فقد قراءة أو أكثر
يعني أن يفقد النموذج توازنه ، ولا تصبح القطاعات متعامدة في أثرها مع
المعالجات ، مما يستوجب تقدير القراءة (أو القراءات) المفقودة وإجراء التحليل

بعد إحلال القراءة المفقودة بتقديرها، وسوف تقتصر على كيفية معالجة تقدير قراءة مفقودة واحدة.

وأحد طرق تقدير القراءة المفقودة ، والتي سنرمز لها بالرمز s^* والتي تجعل m ، للبقاى أصغر ما يمكن هي:

$$s^* = \frac{l \times r^* + r \times r^* - r^* m}{(1-r)(1-l)} \quad (10/2)$$

حيث :

r^* : مجموع المعالجة المحتوية على القراءة المفقودة.

r : مجموع القطع المحتوى على القراءة المفقودة.

m : المجموع الكلى:

ونضع s^* مكان القراءة المفقودة وتحلل نتائج التجربة على هذا الأساس مع تخفيض درجات الحرية للتيارين الكلى والخطأ العشوائى كل بدرجة واحدة.

مثال (٣-٢):

نفرض أن المفردة الأخيرة فى العمود الأول من جدول مثال (٣-٢) وهى تمثل قيمة الوحدة التجريبية بعد المعالجة الأولى فى اليوم السادس قد فقدت (ليس نسيجة للتجربة وإلا كانت = صفر) . أوجد تقدير هذه القيمة واستكمل التحليل.

$$r^* - r = 1479 - 190 = 1289$$

* لتقدير أكثر من قراءة مفقودة يرجع إلى بعض المراجع الأكثر تخصصاً مثل:

Steel, R.G.D., Torrie, J.H.; Principles And Procedures of Statistics, McGraw - Hill Book company 1960.

$$١٠٣٢ - ١٩٥ - ٨٣٧ = ١٠٣٢$$

$$٨٠٩٧ - ١٩٥ - ٨٢٩٢ = ٨٠٩٧$$

$$\frac{٨٠٩٧ - ١٢٨٤ \times ٦ + ٨٣٧ \times ٨}{٧ \times ٥} = ١٨٠$$

$$١٨٠ = \frac{٨٠٩٧ - ٧٧٠٤ + ٦٦٩٦}{٣٥}$$

أما مجاميع المركبات التي حسبت على أساس س* فهي كالآتي:

$$١٦٧ + ٠٠٠ + ١٧٢ + ١٨٠ + ٠٠٠ + ١٧٧ + ١٦٤ = ٨٢٧٧$$

$$٨٦٩٦ = ١٤٢٧٢٦٥ - ١٤٣٥٩٦١ = ٨٦٩٦$$

م.م. بين الدهون (معالجات)

$$\frac{٨٢٧٧}{٤٨} - \left(\frac{١}{١} (١٠١٧ + ١٠٦٧ + ٠٠٠ + ٩٧٤) \right) = ٣٣٩٩,٨$$

$$٣٣٩٩,٨ = ١٤٢٧٢٦٥ - ١٤٣٠٦٦٤,٨ = ٣٣٩٩,٨$$

م.م. بين الأيام (قطاعات)

$$\frac{٨٢٧٧}{٤٨} - \left(\frac{١}{٨} (١٤٦٤ + ٠٠٠ + ١٤٧٢ + ١٣٣١) \right) = ٣٦٤٦,٦$$

$$٣٦٤٦,٦ = ١٤٢٧٢٦٥ - ١٤٣٠٩١١,٩ = ٣٦٤٦,٦$$

أن س.م. البواقي = ١٦٤٩,٦

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
الأيام (قطاعات)	٣٦٤٦,٦	٥	٧٢٩,٣٢	١٥,٠٣
الدهون (معالجات)	٣٣٩٩,٨	٧	٤٨٥,٦٩	١٠,٠١
البواقي	١٦٤٩,٦	٣٤	٤٨,٥٢	
كلى	٨٦٩٦,٠	٤٦		

ولازال فرضى العدم مرفوضين عند $\alpha = 1\%$ حيث:

$$F_{1, 34, 0} = 3,70 > F_{1, 34, 0,7} = 3,30$$

(١-٥) : الخطأ المعياري للفرق بين أى متوسطين:

والخطأ المعياري للفرق بين أى متوسطين هو كالمعتاد

$$E_{\bar{X} - \bar{Y}} = \sqrt{\left(\frac{1}{n} + \frac{1}{n}\right) \sigma^2} = \sqrt{\left(\frac{2}{n}\right) \sigma^2}$$

حيث $n =$ ر إذا كان الخطأ المعياري بحسب للفرق بين متوسط أى معالجتين أو
 $n =$ ل إذا كان بحسب للفرق بين متوسط أى قطاعين. (١، ٢) فى دليل \bar{S}
 تشير إلى المعالجتين أو القطاعين موضوع المقارنة ، \bar{X} هى متوسط مربعات
 البواقي فى جدول تحليل التباين.

وبالتالى فإن الحد الأدنى للفرق المعنوية لأى متوسطين * هو:

$$E_{\bar{X} - \bar{Y}} = \sqrt{\left(\frac{2}{n}\right) \sigma^2} = \sqrt{\left(\frac{2}{n}\right) \sigma^2}$$

وعليه فإن الخطأ المعياري للفرق بين أى معالجتين من بيانات المثال (٢ - ٢)
 وجدول تحليل التباين له هو:

$$E_{\bar{X} - \bar{Y}} = \sqrt{\left(\frac{2}{n}\right) \sigma^2} = \sqrt{\left(\frac{2}{n}\right) \sigma^2}$$

والحد الأدنى للفرق المعنوية = $4,104 \times 3,30 = 13,92$
 وبالنسبة للقطاعات:

$$E_{\bar{X} - \bar{Y}} = \sqrt{\left(\frac{2}{n}\right) \sigma^2} = \sqrt{\left(\frac{2}{n}\right) \sigma^2}$$

* حسب هذه الفترة على أساس متوسط فترة الثقة للفرق بين أى متوسطين يختارا مقبلا إجراء التجربة وإلا
 كانت الفترة أكثر اتساعاً. ولمعالجة هذه النقطة يرجع إلى:

Anderson, R.L. Bancroft; T.A. "Statistical Theory in Research, McGraw - Hill Book Company 1952, P. 239.

والحد الأدنى للفروق المعنوية $= |3,60 \times 3,35| - |12,06|$
 وفي حالة وجود قراءة مفقودة فإن الخطأ المعياري للفرق بين متوسط
 المعالجة ذات القراءة المفقودة ومتوسط أى معالجة أخرى يكون:

$$E_u = \sqrt{\left(\frac{L}{(1-J)(1-R)} + \frac{2}{J} \right) \sigma^2}$$

حيث ع ٢ : متوسط مربعات البواقي من جدول تحليل التباين المحسوب بعد
 تقدير القراءة المفقودة.

وتطبيقاً لذلك على المثال (٣ - ٢) فإن:

$$E_u = \sqrt{\left(\frac{8}{7 \times 5 \times 6} + \frac{2}{6} \right) 48,52} = 4,24$$

$$4,24 = (0,038 + 0,333) 48,52$$

(٣) القطاعات الكاملة العشوائية مع التكرار:

(٣-١) مقدمة:

كثيراً ما تكون الملاحظة ليست هى وحدة التجربة ، ولكن الوحدة تشتمل
 على عدة مفردات ، يسجل عن كل مشاهدة نتيجة إجراء التجربة، وحين تتعدد
 المشاهدات لكل وحدة تجربة أى يكون هناك عدة مشاهدات لكل قطاع / معالجة
 فإنه يتعين التمييز بين خطأ التجربة (Experimental Error) والخطأ العشوائى
 (Random error or Residual) وسوف نقتصر على معالجة حالة تساوى عدد
 التكرارات (replicates) لكل وحدة تجربة ، أى لكل معالجة وقطاع. وفي هذه
 الحالة فإن النموذج الرياضى يكون كما يلى:

(٣-٢) النموذج الرياضى:

$$Y_{rjk} = \mu + \tau_r + \beta_j + \gamma_k + \epsilon_{rjk}$$

(١٩/٢)

وهذا النموذج لا يختلف في فروضه عن سابقة (٢-١) إلا في إضافة الحد (م-ق) رر ، وهو يعبر عن اختلاف درجة الاستجابة للمعالجات المختلفة من قطاع لآخر . أو بعبارة أخرى فإنه يعنى وجود أثر لكل معالجة / قطاع (معا) على الوحدة التجريبية ، بالإضافة إلى الأثر الثابت لكل معالجة والأثر الثابت لكل قطاع وهو بهذا المفهوم يعرف باسم خطأ التجربة (Experimental Error) ، وهذا الأثر لا يعتبر عشوائياً بالطبع لأنه ينشأ من عوامل معروفة ذات أثر يمكن قياسها وهى اختلاف القطاعات وهى ليست عشوائية. وهذا ما يميز مجموع المربعات لخطأ التجربة عن مجموع المربعات الذى يعزى إلى الخطأ العشوائى أو خطأ العينة (Sampling Error) أو البواقي الذى يقيس اختلاف القراءات أو المشاهدات داخل أى وحدة تجربة عن بعضها البعض.

هذا وقد أخذ فى الاعتبار أن النموذج المتبع فى تخطيط وتنفيذ التجربة هو النموذج الثابت (Fixed Model) أى أن التجربة تشمل جميع المعالجات وجميع القطاعات الممكنتين وليست عينات عشوائية من مجتمعات أكبر وأوسع من المعالجات والقطاعات ، وبالتالي فإن الأثر المضاف لكل معالجة/ قطاع، يعتبر فى حكم التفاعل الثابت (Fixed interaction) . وطالما أن $k > 1$ ، فإنه يمكن استخدام م.م.م. للبواقي أو الخطأ العشوائى لاختبار الفروض الخاصة بالمعالجات والقطاعات وخطأ التجربة ، وهذا ما يميز النموذج الثابت عن النموذج العشوائى (Random Model) والذى تكون فيه المعالجات والقطاعات ليست سوى عينة من مجتمع المعالجات والقطاعات ، وبالتالي فإن الاستنتاج الإحصائى يقتصر على المعالجات والقطاعات التى أجريت عليها التجربة. أو النموذج المختلط (Mixed Model) حين تكون المعالجات أو القطاعات عينة من المجتمع الأكبر ويكون العنصر الآخر أى المعالجات أو القطاعات التى لم تعامل كعينات هى كافة ما هو متاح منها. وفى أن من هاتين الحالتين يكون مقام ف مختلف عن سابقة فى الحالة الأولى. ونكتفى بهذا القدر فى معالجة هذا الأمر.

ويكون جدول تحليل التباين في حالة النموذج الثابت كالتالى:

(٣-٣) نموذج جدول تحليل التباين لتصميم القطاعات الكاملة العشوائية

مع التكرار (النموذج الثابت) :

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية
القطاعات	$\text{مح ر} = \frac{\sum (\sum_{j=1}^r \sum_{k=1}^k x_{jk})^2}{l \times k} - \frac{\sum x_{jk}^2}{r \times l \times k}$	$r - 1$
المعالجات	$\text{مح ر} = \frac{\sum (\sum_{j=1}^r x_{jk})^2}{r \times k} - \frac{\sum x_{jk}^2}{r \times l \times k}$	$l - 1$
خطأ التجربة	$\text{مح ر} = \frac{\sum (\sum_{j=1}^r \sum_{k=1}^k x_{jk})^2}{k} - \frac{\sum x_{jk}^2}{r \times l \times k} - \text{م.م.ق} - \text{م.م.م}$	$(r-1)(l-1)$
البواقي (خطأ عشوائى)	$\text{مح ر} = \left(\sum x_{jk}^2 - \frac{\sum (\sum_{j=1}^r \sum_{k=1}^k x_{jk})^2}{k} \right)$	$rl - (k-1)$
الكلى	$\text{مح ر} = \sum x_{jk}^2 - \frac{\sum x_{jk}^2}{r \times l \times k}$	$rl - k - 1$

وسوف نستخدم بيانات المثال التالى (٢-٤) كشرح عملى لأسلوب

التحليل فى هذه الحالة .

مثال (٣-٤) :

بفرض أن البيانات التالية تمثل إنتاج ثلاثة أنواع من الآلات (ل = ٣ معالجات) فسي خمسة مواقع إنتاجية (ر = ٥ قطاعات) حيث سجلت قراءتين فسي كل موقع على كل آلة (أي أن الوحدة التجريبية تتكون من ك = ٢ مشاهدة).

م ل	مواقع الإنتاج (القطاعات)					أنواع الآلات (معالجات)
	(٥)	(٤)	(٣)	(٢)	(١)	
٢٤٥	٢٦	٢٨	٢٩	٢٨	٢٥	أ
	٢٠	٢٢	٢٠	٢٥	٢٢	
٢٨٠	٢٧	٣١	٣٠	٢٤	٢٥	ب
	٢٥	٣٠	٢٨	٣٠	٣٠	
٢٥٥	٢٠	٢٣	٢٥	٣٠	٢٥	ج
	٢٩	٢٨	٢٦	٢٠	٢٩	
٧٨٠	١٤٧	١٦٢	١٥٨	١٥٧	١٥٦	د

$$\frac{780}{30} - 29 + \dots + 28 + 25 = \text{مجموع المربعات الكلية}$$

$$344 = 20280 - 20624 =$$

مجموع المربعات بين القطاعات

$$\frac{780}{30} - \left(\frac{147 + 162 + 158 + 157 + 156}{3 \times 2} \right) =$$

$$20,33 =$$

مجموع المربعات بين المعالجات

$$60 = \frac{780}{30} - \left(\frac{250 + 280 + 245}{5 \times 2} \right) =$$

مجموع المربعات للتفاعلات (أو خطأ التجربة)

$$-m \cdot m \cdot m - q \cdot m \cdot m - \frac{2780}{5 \times 3 \times 2} - (29 + \dots + 53 + 27) \cdot \frac{1}{2} =$$

= مجموع المربعات بين الخلايا (وحدات التجربة) - م.م.ق - م.م.م - م

$$27,77 = 2.,33 - 70 - (2.28. - 2.392) =$$

مجموع المربعات للخطأ العشوائي (أو البواقي)

= م.م. الكلى - م.م. بين الخلايا.

$$232 = 122 - 366 =$$

$$10 = 14 - 29 = (8+2+4) - 29 = \text{بدرجات حرية}$$

ويكون جدول تحليل التباين كالآتي:

فـ α - %	نسبة التباين فـ *	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر الاختلافات
٣,٠٦	٠,٣٣	٥,٠٨	٤	٢٠,٣٣	بين القطاعات
٣,٦٨	٢,١٠	٣٢,٥٠	٢	٦٥	بين المعالجات
٢,٦٤	٠,٢٢	٣,٣٣	٨	٢٦,٦٧	خطأ التجربة
			١٤	١١٢	بين وحدات التجربة (الخلايا)
	١٥,٤٧		١٥	٢٣٢	البواقي (الخطأ العشوائي)
	—	—	٢٩	٣٤٤	الكلي

* بالفرق ، ** بالفرق

وتقبل الفروض بأن $\frac{1}{2} = \frac{1}{3} = \frac{1}{4} = \text{صفر}$

ق_۱ = ق_۲ = ... = ق_۰ = صفر

$$(ق-1) = \dots = \dots = (ق-2) = (ق-1) - 1 = \text{صفر}$$

عند مستوى المعنوية ٥ %.

ثانياً : المربع اللاتيني

LATIN SQUARE

(١) مقدمة:

وفي هذا النموذج نوزع كل معالجة عشوائياً داخل كل صف وكل عمود، بحيث تظهر كل المعالجات في كل الصفوف وكل الأعمدة بشرط أن تظهر كل معالجة مرة واحدة وواحدة فقط داخل كل صف وكل عمود وبالتالي يمكن إزالة كل تباين نتيجة اختلاف الأعمدة والصفوف من الخطأ. وفي هذا النموذج يكون عدد الصفوف = عدد الأعمدة = عدد المعالجات = r مثلاً. ففي تجربة تسويقية يمكن أن تكون الأيام هي الصفوف والمتاجر هي الأعمدة وأنواع السلع هي المعالجات كما أنه في تجربة أخرى يمكن أن تقوم الأيام مقام القطاعات (٤ أيام أسبوعياً مثلاً) وهي الأعمدة. كما أن الصفوف (٤) تمثل أنواع مختلفة من الآلات والورديات هي العامل الثالث (المعالجات وعددها ٤ أيضاً) ولنموذج المربع اللاتيني تطبيقات عديدة في مجال التجارب الزراعية.

ويمكن أن نصل إلى عدد كبير من التوليفات الممكنة لكل عدد من المعالجات ، ففي حالة مربع لاتيني 4×4 فيوجد ٥٧٦ توليفة ممكنة أحداها النموذج التالي:

المعالجات (الأعمدة)				
III	II	I		
د	جـ	ب	أ	(١)
جـ	ب	أ	د	(٢) قطاعات
ب	أ	د	جـ	(٣)
أ	د	جـ	ب	(٤)

حيث تمثل أ ، ب ، جـ ، د الأنواع الأربعة للعامل الثالث الذي ادخل على الأعمدة والصفوف.

وفى حالة مربع لاتينى 5×5 فإنه يوجد ١٦١٢٨٠ توليفة ممكنة ،
يمكن اختيار أحدها بالرجوع إلى أحد المراجع المتخصصة الواردة فى قائمة
المراجع فى نهاية هذا المرجع.
والعدد المناسب لاستخدام هذا النموذج يتراوح بين 5×5 إلى 8×8 إذا
أنه مع زيادة العدد عن ذلك ، سيصبح عدد الوحدات التجريبية غير عملى . ومع
ذلك فإن عدد درجات حرية الخطأ لهذا النموذج = $n^2 - 1$ أى أنه عند $n = 3$
فإن درجات حرية الخطأ = ٨ ويعنى ذلك ضرورة مراعاة اختيار العدد المناسب
من الناحية العملية ولتوفير العدد المناسب لدرجات حرية الخطأ.
وسوف نستخدم بيانات المثال التالى (٢ - ٥) لعرض نتائج تجربة مربع لاتينى
 4×4 .

مثال (٢ - ٥):

البيانات التالية تمثل نتائج تجربة مربع لاتينى 4×4 :

مواقع الإنتاج (قطاعات)					أنواع الآلات (معالجات)
م	III	II	I	م	
٢١٩	د/٥٣	ج/٤٤	ب/٤١	أ/٨١	(١)
٢٢٦	ج/٤٩	ب/٤٢	أ/٩٧	د/٣٨	(٢)
١٧٧	ب/٣٦	أ/٦٧	د/٤٣	ج/٣١	(٣)
٢١٤	أ/٨١	د/٤٣	ج/٣٣	ب/٥٧	(٤)
م = ٨٣٦	٢١٩	٤٩٦	٢١٤	٢٠٧	م
الأيام	د	ج	ب	أ	م
	١٧٧	١٥٧	١٧٦	٣٢٦	
م = ٨٣٦					

وتحليل النتائج يبين:

$$\text{مجموع المربعات الكلى} = ٨١^2 + ٤١^2 + ٥٠^2 + ٨١^2 - \frac{٨٣٦^2}{4 \times 4} = ٥٦٦٧$$

$$\text{مجموع المربعات للأعمدة (مواقع الإنتاج)} = \frac{1}{4} - \frac{836}{4 \times 4} = \frac{214 + 207}{4} - \frac{836}{16} = 74,5 - \frac{836}{16} = 74,5 - 52,25 = 22,25$$

$$\text{مجموع المربعات للصفوف (الآلات)} = \frac{1}{4} - \frac{836}{4 \times 4} = \frac{226 + 219}{4} - \frac{836}{16} = 359,5 - \frac{836}{16} = 359,5 - 52,25 = 307,25$$

$$\text{مجموع المربعات لأيام} = \frac{1}{4} - \frac{836}{4 \times 4} = \frac{217 + 157 + 217 + 226}{4} - \frac{836}{16} = 462,5 - \frac{836}{16} = 462,5 - 52,25 = 410,25$$

ويكون مجموع المربعات المتبقى = 5667 - 74,5 - 307,25 - 410,25 = 5255

$$= 606,5$$

جدول تحليل التباين

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
الأعمدة (مواقع الإنتاج)	74,5	3	24,8	0,25
الصفوف (آلات)	307,25	3	102,4	1,18
الأيام	462,5	3	154,2	15,25
البواقي	606,5	6	101,1	—
الكلي	5667,0	15	—	—

ولكن ف ٠,٠٠٢ % = ٤,٧٦ لذلك يقبل الفرض بعدم وجود أثر لكل من الأعمدة والصفوف ، ويرفض الفرض العدمي الخاص بالأيام بمعنى أن الأيام ذات أثر معنوي على وحدات التجربة.

هذا ويلاحظ أن هذا النموذج يختلف عن التحليل في اتجاهين ، في أنه يفترض عدم وجود تفاعل بين الصفوف والأعمدة ، كما أن عدد وحدات التجربة = ٢ × ٢ أي ٤ رغم دخول ثلاثة عوامل في التجربة مما كان يعني استخدام ٢ × ٢ × ٢ وحدة تجريبية.

(٢) النموذج الرياضي والجبري (المسابى):

وعموما فإن نموذج المربع اللاتيني هو:

س ر ل ك = م + ن + ق + ع + ح ر ل ك

ر، ل، ك = ١، ٢،، ن

أما فروض النموذج فهي نفس الفروض السابقة لعامل وعاملين ، مع إضافة العامل الثالث ع وهي نموذج الآثار المضافة ، أى أن:

عج رل (س رل - س ...)^۲ = ن عج (س رل - س ...)^۲

$$+n \text{ مَجْ } (\text{م.} - \text{ج.} \dots) + n \text{ مَجْ } (\text{م.} - \text{ك.} \dots)$$

ج + (من روك - .. ر - .. ر - .. ر + ... ر) (٢١/٢)

أو

م.م.ك = م.م. بين الأعمدة + م.م. بين الصفوف + م.م. بين المعالجات

+ م.م. للخطأ العشوائي (أو البواقى) (٢٢/٢)

وتحسب مجاميع المربعات كالسابق.

كما أن الخطأ مستقل عن المعالجات والقطاعات والعوامل ، وأن لها توزيع معتاد (صفر ، σ) وبالتالي لم يحدث اختلاف في أسلوب التحليل عما سبق وهو ما يتبين لنا من المثال (٢ - ٥) .

هذا ويتم العشوائية في نموذج المربع اللاتيني باختيار توليفة من كافة التوليفات الممكنة بطريقة عشوائية*.

* يمكن الرجوع إلى جدول فيشر وتبني للحصول على الصور الممكنة للمربع اللاتيني ٦×٦ حتى ٢٣٢
Fisher, R.A. and Yates, F., Statistical Tables for Biological, Agricultural And Medical Research, 5th ed., Hafner publishing company, N.Y., 1957.

(٣) الكفاءة النسبية لنموذج المربع اللاتيني:

لقياس كفاءة المربع اللاتيني مقارنة بالقطاعات الكاملة العشوائية يستخدم المعامل التالي:

$$م.ك.ن(م لاتيني / ق ع) = \frac{ع^2}{ع} \times \text{معامل التصحيح} = م.م.ق. + \frac{ع^2(١-ن)}{ن(١-ع)} \times \frac{(١+٢ن)(٣+١ن)}{(١+١ن)(٣+٢ن)} \times ١٠٠ \times (٢٣/٢)$$

حيث:

ع^٢: متوسط مربعات البواقي للمربع اللاتيني.

م.م.ق: مجموع مربعات القطاعات.

(ن-١): درجات الحرية للقطاعات (أو المعالجات لانهما متساويان)

ن: درجات حرية ع^٢ للمربع اللاتيني

١: درجات حرية ع^٢ للقطاعات الكاملة العشوائية.

وجدير بالذكر أننا نقدر كفاءة تصميم المربع اللاتيني من واقع ناتج

تحليله مقارنة بالقطاعات العشوائية الكاملة تقديرًا للخطأ العشوائي لو أننا استبدلنا

من البداية نموذج المربع اللاتيني بالقطاعات الكاملة العشوائية وذلك يعتمد على

مقارنة ع^٢ للأخير مقدرة من ع^٢ لتصميم المربع اللاتيني كما سبق أن بينا في

حالة القطاعات الكاملة العشوائية ، أما الحد الأخير في (٢٣/٢) فهو تصحيح

لنقص درجات الحرية باستخدام المربع اللاتيني.

وتطبيقاً لذلك على بيانات المثال (٥-٢) نجد الآتي:

$$م.ك.ن = \frac{١٠١,١ \times ٩ + ٤٧,٥}{١٠١,١ \times ٣ \times ٤} \times \frac{(٣+٩)(١+٦)}{(١+٩)(٣+٦)} \times ١٠٠ \times (٢٣/٢)$$

$$م.ك.ن = \frac{٩٠٩,٩ + ٤٧,٥}{١٠١,١ \times ١٢} \times \frac{١٢ \times ٧}{١٠ \times ٩} \times ١٠٠ \times (٢٣/٢) = ٧٥,٧ \%$$

أى أن القطاعات الكاملة العشوائية أعلى كفاءة من المربع اللاتينى فى حالتنا هذه (باستخدام الأعمدة كقطاعات) وأنه يمكن الوصول إلى نفس الدرجة من الكفاءة (نفس الخطأ العشوائى) بـ ٧٥ مفردة باستخدام القطاعات الكاملة العشوائية مقابل ١٠٠ مفردة باستخدام المربع اللاتينى.
وإذا اعتبرنا الصفوف هى القطاعات ، فإن كفاءة النموذج مقارنا بالقطاعات الكاملة العشوائية هى:

م.ك.ن (مربع لاتينى / ق / ع (للفوف)

$$\%97,7 = 100 \times \frac{84}{90} \times \frac{101,1 \times 9 + 309,5}{101,1 \times 3 \times 4} =$$

وباعتبار الصفوف هى القطاعات ، فإن كفاءة المربع اللاتينى تقترب من كفاءة القطاعات الكاملة العشوائية. وهكذا تزداد كفاءة استخدام المربع اللاتينى كلما كانت القطاعات متباينة الأثر فيما بينها.

(٤) القراءة المفقودة:

وبفرض أن أحد القطع أو أحد الوحدات التجريبية فقدت لسبب آخر لا يرجع إلى التجربة ذاتها ، فإنه يمكن تقدير قيمة هذه القراءة المفقودة بنفس الأسلوب الذى اتبع فى السابق ويكون تقديرها كالاتى:

$$\text{ص} = \frac{\text{ن}(\text{م}^* + \text{م}^* + \text{م}^* + \text{م}^* - \text{م}^*)}{(\text{ن} - 1)(\text{ن} - 2)} \quad (24/2)$$

حيث م^{*}، م^{*}، م^{*}، م^{*}، م^{*} هى مجموع الصف والعمود والمعامل والكلى الذى يشتمل على القراءة المفقودة.

مثال (٣-٦):

وبفرض أن المفردة (I ، ٤ ، ب) أى المفردة فى نهاية العمود الأول وبداية الصف الرابع قد فقدت قيمتها الحقيقية فإن تقديرها يكون كالاتى:

$$م^* = ١٥٧ ، م^* = ١٥٠ ، م^* = ١١٩ ، م^* = ٧٧٩$$

$$\text{أذن ص}^* = \frac{٧٧٩ \times ٢ - (١١٩ + ١٥٠ + ١٥٧)٤}{٢ \times ٣} = ٢٤,٣$$

وتستبدل القراءة الأخيرة المفقودة بالقيمة المقدرة ٢٤,٣ ويعاد التحليل من جديد، وتنقص درجات حرية المجموع الكلي والبقوى كل بدرجة واحدة مقابل تقدير القراءة المفقودة.

أما الخطأ المعياري للفرق بين متوسطى أى معالجتين تحتوى أحدهما على القراءة المفقودة فهو:

$$ع_١ - ع_٢ = \sqrt{\frac{١}{(٢-ن)} + \frac{٢}{ن}} \cdot ٢٤,٣$$

و $ع_١ \neq ع_٢$ ، $٢٤,٣$ = متوسط مربعات البواقى من جدول تحليل التباين.

ثالثاً: التحليل العاملي

Factorial Analysis

(1) مقدمة: العامل أو الأثر الأساسي والتفاعل:

التحليل العاملي (Factorial Analysis) هو تحليل نتائج تجارب تصميم خصيصاً بفرض إتاحة الفرصة لإجراء مقارنات مستقلة تعتمد على طريقة اختيار المعالجات ، والتي تمثلها في هذا الأسلوب العوامل الرئيسية والتفاعلات ، وذلك باستخدام أسلوب تحليل التباين ، وكل عامل يتكون من عدة مستويات (Levels) ، ويعتبر أن نعين الأثر الأساسي المضاف لكل عامل عند مستوياته الممكنة ، وكذلك لكل مستوى للعامل مع المستويات المختلفة لعامل آخر ، ويعرف باسم التفاعل (Interaction) فمثلاً إذا اعتبر الحافز النقدي عاملاً له أثره على الإنتاج ، فإنه يمكن أن يتكون من عدة مستويات، وكل مستوى يحدد -مثلاً- على أساس نسبة معينة من المرتب أو الأجر الشهري، كما قد يعرف على أساس مبلغ معين من المال يصرف كحافز سنوي. وبالمثل إذا اعتبر البرنامج التدريبي عاملاً ، فإنه يمكن أن يتكون من عدة مستويات، فقد يكون البرنامج على مستوى المقدمة ثم برنامج متوسط المستوى وثالث متقدم.... وهكذا.

ولتصميم تجربة كهذه تشتمل على عامل واحد (Single Factor) مثل الحافز النقدي وأثره على الإنتاج ، وبفرض أن هذا العامل يتحدد بثلاثة مستويات هي ١٠% ، ٢٠% ، ٣٠% من الأجر الشهري للعامل ، فإن ذلك يتطلب المحافظة على بقية المتغيرات - عدا الحافز النقدي - متجانسة أو ثابتة قدر الإمكان. وإذا ما أدخل عامل ثان في الاعتبار كمدة الخبرة في ممارسة العمل ، مثلاً ، وبفرض أنها قسمت إلى ثلاثة فترات (أو مستويات) هي أقل من ٥ سنوات ، من ٥ سنوات إلى أقل من ١٠ سنوات ، ثم ١٠ سنوات فأكثر، فإن ذلك

يعنى أننا بصدد تجربة مزدوجة العوامل، تتكون من 3×3 أو 2×3 - تسعة صور ممكنة لمستويات العامل الأول مع مستويات العامل الثانى حيث يظهر كل مستوى من الحافز النقدي فى فترات الخبرة المختلفة ويمكن تصوير ناتج هذه التجربة بالجدول التالى:

ب: مدة الخبرة بالسنوات	أ : الحافز النقدي		
	أ ₁ %١٠	أ ₂ %٢٠	أ ₃ %٣٠
ب ₁ : ٥-	أ ₁ ب ₁	أ ₂ ب ₁	أ ₃ ب ₁
ب ₂ : ٥-	أ ₁ ب ₂	أ ₂ ب ₂	أ ₃ ب ₂
ب ₃ : ١٠ فأكثر	أ ₁ ب ₃	أ ₂ ب ₃	أ ₃ ب ₃

وقد استخدمت الحروف الأبجدية أ ، ب ، ... للدلالة على العوامل ويشير الدليل أدنى كل حرف إلى المستوى الخاص بالعامل ، فقد استخدم الحرف أ للدلالة على العامل الأول وهو الحافز النقدي ، ب للعامل الثانى، وهو طول مدة الخبرة ، أما أ₁ فتشير إلى المستوى الأول أو الأدنى من العامل الأول أ ، أ₂ إلى المستوى الثانى لنفس العامل ... وكذلك ب₁ ترمز إلى أدنى مستوى للعامل الثانى ب ، وبالمثل ب₂ ترمز إلى المستوى الثانى لهذا العامل ، أما أ₃ ب₃ (مثلاً) فتشير إلى أثر المستوى الثانى من العامل الأول أ مع المستوى الثانى للعامل ب أو المعالجة المشتركة للعاملين عند هذين المستويين، ويشار إليه فى التحليل العاملى بأنه تفاعل أ₃ مع ب₃ (Interaction) وهو يشير إلى التفاوت أو التباين فى أثر أ عند المستويات المختلفة لب ، وهو يعنى أن هذا التباين لا يرجع إلى الصدفة أو العشوائية.

ويقوم أسلوب التحليل العامل على أساس اعتبار المعالجات كمعامل ذات مستويات مختلفة. وباستخدام أسلوب تحليل التباين، يمكن تجزئة مجموع المربعات للمعالجات إلى مكونات بعضها يعبر عن الأثر الأساسي (Main effect) لكل عامل أ، ب، ج، والبعض يعبر عن الأثر المشترك لعاملين أو أكثر أى التفاعلات، ويمكننا ذلك من تحديد أفضل المستويات للعوامل فى إحداث أكبر أثر معنوى على وحدات التجربة.

وجدير بالذكر أنه حتى لو حدث وتبين عدم معنوية التفاعلات، فإن ذلك لا يقلل من نتيجة التجربة أو إلى أهدار معلومات، بل أن ذلك يؤدي إلى تكرار أكبر للوحدات التجريبية لا يؤثر على دقة النتائج وإنما يعتبر تكرار مقنعا. وحين يتبين أن التفاعلات معنوية، فإن ذلك يشير إلى عدم استقلال أثر العوامل. وقد عرفت هذه النماذج فى البداية فى التجارب الزراعية، ثم عرفت طريقها بعد ذلك فى مجالات أخرى متعددة خاصة حين يكون الهدف تحديد المستوى الأمثل لعامل ما مع مستويات عامل آخر، كتحديد درجة الحرارة المناسبة مع الفترة الزمنية للوصول إلى درجة صلابة معينة للنتاج أو أى تركيز مناسب لنوع من السماد مع أى نوع من البذور للوصول إلى أفضل محصول... وهكذا.

وسوف نقتصر فى معالجة النماذج التى سوف نقدمها عن التحليل العامل فى الفقرات التالية، على تخصيص عد متساو من الوحدات التجريبية لكل مستوى من العامل مع أى مستوى لعامل آخر مقتصرين فى ذلك على التجارب 2×2 أو 2^2 أى لعاملين لكل مستويين.

(٣) التحليل العامل لتجربة 2×2 :

وفى هذه الحالة سوف يكون لدينا عاملين أ، ب كلاهما له مستويين وبالتالي فإن المعالجات سوف تنقسم إلى الأثر الرئيسى لكل من أ، ب ثم الأثر المشترك (التفاعل) لـ أ ب وتعبير عنه المجموعات أ_١ ب_١، أ_١ ب_٢، أ_٢ ب_١، أ_٢ ب_٢

وسوف يكون النموذج الرياضى فى حالة استخدام أسلوب التحليل فى اتجاه واحد أو النموذج العشوائى الكامل كالاتى:

(١-٣) النموذج الرياضى:

س ر د ر = $\mu + \alpha + \beta + \gamma + (\alpha\beta) + \gamma\delta + \epsilon$ (٢٦/٢)
حيث ترمز α إلى الأثر المضاف للعامل α ، (ل = ١ ، ٢) ، β إلى الأثر المضاف للعامل الثانى β (ك = ١ ، ٢) ، γ ترمز إلى عدد وحدات التجربة لكل خلية (α ، β ، γ) ، أما $\gamma\delta$ فتشير إلى التباين العشوائى ، وسوف نفترض أن النموذج لاأثر مضافة وأن $\gamma\delta$ مستقلة عن α ، β وأن لها توزيع معتاد توقعه الصفر وتباينه σ^2 كالمعتاد . كما أن ($\alpha\beta$) $\gamma\delta$ تشير إلى تفاعل α مع β أو اختلاف أثر α مع مستويات β ، وأن :
 $\epsilon = (\alpha) - \mu - \alpha - \beta - \gamma - (\alpha\beta) - \gamma\delta - \epsilon$ صفر .

وسوف نقصر فى معالجة حساب مجموع المربعات للمعالجات وتجزئته إلى مجموع المربعات للعوامل وللتفاعلات بنفس الأسلوب الذى اتبع فى السابق وأن كانت هناك طريقة أخرى تعتمد على ما يعرف باسم المعاملات الخطية المتعامدة أو كثيرات الحدود المتعامدة (Orthogonal Polynomials or Orthogonal Linear Forms) ولكننا سوف لا نتعرض لهذه الطريقة الأخيرة فى الوقت الحاضر .

مثال (٧-٣) :

لدراسة أثر العامل (أ) على تركيز مادة ما عند مستويات مختلفة من العوامل (ب) وكان كل من العاملين يتميز بمستويين فقط، فقد خصص ر = ٥ وحدات تجريبية عشوائيا للتجميعات الأربعة للعاملين α_1 ، α_2 ، β_1 ، β_2 ، وكانت النتائج كالاتى:

* Steel m R.G.D., Torrie, J.H., : Principles and Procedures of Statistics, McGraw-Hill Book Company Inc., 1960, PP 199 – 202.

	أ ₁	أ ₂	أ ₃	أ ₄	
	أ ₁ ب ₁	أ ₁ ب ₂	أ ₁ ب ₃	أ ₁ ب ₄	
	٨,٥٣	١٧,٥٣	٣٩,١٤	٣٢,٠٠	
	٢٠,٥٣	٢١,٠٧	٢٦,٢٠	٢٣,٨٠	
	١٢,٥٣	٢٠,٨٠	٣١,٣٣	٢٨,٨٧	
	١٤,٠٠	١٧,٣٣	٤٥,٨٠	٢٥,٠٦	
	١٠,٨٠	٢٠,٠٧	٤٠,٢٠	٢٩,٣٣	
مجموع	٦٦,٣٩	٩٦,٨٠	١٨٢,٦٧	١٣٩,٠٦	٤٨٤,٩٢
مجموع	٩٦٣,٨٨	١٨٨٧,٠٢	٦٩١٣,٦٣	٣٩١٢,١٢	١٣٦٧٦,٧٠

ولتحليل نتائج هذه التجربة فإن ذلك يمكن أن يتم على خطوتين:

الخطوة الأولى:

وتحسب فيها مجموع المربعات المعتاد لتحليل التباين وهي:

مجموع المربعات الكلى = م. م. ك = مجموع^٢ روك - $\frac{\sum (م)^2}{رل ك}$

$$= ١٣٦٧٦,٧٠ - \frac{(٤٨٤,٩٢)^2}{٥ \times ٢ \times ٢} = ١٩١٩,٣٣$$

مجموع المربعات للمعالجات =

$$= \frac{(١٣٩,٠٦ + ١٨٢,٦٧ + ٩٦,٨٠ + ٦٦,٣٩)^2}{٥}$$

$$= \frac{(٤٨٤,٩٢)^2}{٥ \times ٢ \times ٢} = ١٥٣٩,٤١$$

مجموع المربعات المتبقى = ١٩١٩,٣٣ - ١٥٣٩,٤١ = ٣٧٩,٩٢

* ونحن هنا بصدد ٤ معالجات يرمز لها بـ أ₁ب₁ ، أ₁ب₂ ، أ₁ب₃ ، أ₁ب₄ ، خصصن لكل عدد ٥ وحدات تجريبية. وبالتالي فإن م.م. للمعالجات يتكون من ثلاثة مركبات هي م.م. أ ، م.م. ب ، م.م. (أب).

الخطوة الثانية:

وفيها يتم تجزئة مجموع المربعات للمعالجات إلى مجموع مربعات لكل عامل أساسي وآخر لكل من التفاعلات كالتالي:

$$\text{مجموع المربعات للعامل (أ) أي م.م. (أ)} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

$$\text{واختصاراً للرموز} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

حيث: أ = ٢ = ١، ٢ = ١، ٢ = ١، ٢ = ١

$$\text{م.م. (أ)} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

$$\frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} = \frac{(484,92)}{5 \times 2 \times 2}$$

$$1256,75 - 11757,37 - \frac{1321,73 + 1163,19}{10} =$$

$$\text{مجموع المربعات للعامل (ب) أي م.م. (ب)} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

$$\text{واختصاراً للرموز} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

* حيث أن ب = ٢ = ١، ٢ = ١، ٢ = ١، ٢ = ١

$$\text{م.م. (ب)} = \frac{\sum_{i=1}^r \sum_{j=1}^l \sum_{k=1}^c (i,j,k)^2}{r \times l \times c} - \frac{\sum_{i=1}^r \sum_{j=1}^l (i,j)^2}{r \times l} - \frac{\sum_{i=1}^r \sum_{k=1}^c (i,k)^2}{r \times c} - \frac{\sum_{j=1}^l \sum_{k=1}^c (j,k)^2}{l \times c} + \frac{\sum_{i=1}^r (i)^2}{r} + \frac{\sum_{j=1}^l (j)^2}{l} + \frac{\sum_{k=1}^c (k)^2}{c}$$

** رمزنا إلى مستويات كل من العامل (أ) والعامل (ب) وعدد كل ٢ بالرمز ل ولو اختلف عدد مستويات أ عن ب لاضطررنا لاستخدام رمز آخر مثل ل للدلالة على مستويات أ ، و (مثلاً) للدلالة على مستويات ب.

$$8,71 = 11707,37 - \frac{235,86 + 249,06}{10}$$

مجموع المربعات للتفاعلات (أ ب) أي م.م. (أ ب) =

$$= \text{مجموع (ب)} - \frac{\text{م}}{\text{ر} \times \text{ل} \times \text{ك}} - \text{م.م. (أ)} - \text{م.م. (ب)} \quad (29/2)$$

$$= 266,39 + 296,80 + 182,67 + 139,06$$

$$8,71 = 1256,75 - \frac{(484,92)}{5 \times 2 \times 2}$$

$$= \text{مجموع المربعات للمعالجات} - \text{م.م. (أ)} - \text{م.م. (ب)} \quad (30/2)$$

$$= 273,95 = 8,71 - 1256,75 - 1539,41$$

ثم يتم تصوير النتائج في جدول تحليل التباين كالتالي:

المصدر للتباين	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
العامل (أ)	1256,75	ر-1 = 1	1256,75	52,92
العامل (ب)	8,71	ل-1 = 1	8,71	0,37
التفاعل (أب)	273,95	(ر-1)(ل-1) = 1	273,95	11,53
المعالجات	1539,41	رل-1 = 3	—	—
البواقي	379,92	رل (ك-1) = 16	23,75	—
الكل	1919,33	رل ك-1 = 19	—	—

* أو بالفرق

ويتبين من التحليل أن العامل (أ) والتفاعلات (أب) كلاهما له أثره

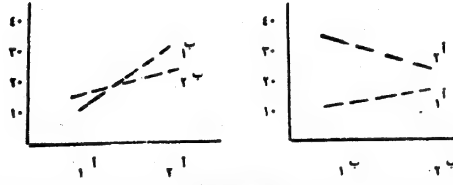
المعنوي عند مستوى المعنوية 1% ، حيث (ف) 11,16 ، حيث (ف) 11,16 - 8,02.

ومعنوية أثر التفاعلات تعني أن العاملين أ ، ب غير مستقلين الأثر ، أي أن الفرق بين الأثر الأساسي للعامل أ عند المستويين للعامل ب أي عند ب_١ ، ب_٢ معنوي ، والعكس أيضاً صحيح بمعنى أن الفرق بين الأثر الأساسي للعامل ب عند المستويين أ_١ ، أ_٢ معنوي.

وإذا رجعنا إلى متوسطات المعالجات لوجدنا أن:

أ _١ ب _١	أ _١ ب _٢	أ _٢ ب _١	أ _٢ ب _٢	م
١٣,٢٨	١٩,٣٦	٣٦,٥٣	٢٧,٨١	

أي أن متوسط الأثر للعامل أ عند المستوى ب_١ قد ارتفع من ١٣,٢٨ إلى ٣٦,٥٣ بالانتقال من المستوى الأدنى لـ أ إلى المستوى الأعلى لذات العامل، أما عند المستوى ب_٢ فقد ارتفع المتوسط من ١٩,٣٦ إلى ٢٧,٨١ . في حين أنه بالنسبة للأثر المتوسط للعامل ب فقد انتقل من ١٣,٢٨ إلى ١٩,٣٦ عند المستوى الأدنى للعامل أ أي أ_١ ، أما عند المستوى الأعلى لـ أ أي أ_٢ فقد انتقل الأثر المتوسط للعامل ب من ٣٦,٥٣ إلى ٢٧,٨١ وهو يبين عدم استقلال أثر كل من العاملين عن العامل الآخر بل ويختلف الأثر اتجاهها وحجمها. ويمكن التعبير عن استجابة كل من العاملين عند كل مستوى للعامل الآخر بالشكل البياني التالي:



ويمكن من دراسة هذا الشكل البياني أو من استكمال تحليل التباين للوصول إلى مجموع مربعات العامل (أ) مع أو من خلال (ب) وكذلك مجموع مربعات (أ) من خلال (ب) ، ومجموع مربعات العامل (ب) من خلال (أ) ، وأخيراً مجموع مربعات العامل (ب) من خلال (أ) وهي:

$$م.م. (أ من خلال ب) = م.م. (أ/ب) = \frac{(م أ/ب - م أ/أ - م ب/أ - م ب/ب)}{ل \times ك}$$

$$(31/2) \quad 1352,10 = \frac{(66,39 - 182,67)}{5 \times 2}$$

$$(32/2) \quad \text{وكذلك م.م. (ب / أ) = } \frac{(م أ/ب - م أ/أ - م ب/أ - م ب/ب)}{ل \times ك}$$

$$178,59 = \frac{(96,80 - 139,06)}{5 \times 2}$$

$$(33/2) \quad \text{أما م.م. (ب / أ) = } \frac{(م ب/أ - م ب/ب - م أ/أ - م أ/ب)}{ر \times ك}$$

$$92,48 = \frac{(66,39 - 96,80)}{10}$$

$$(34/2) \quad \text{وكذلك م.م. (ب / أ) = } \frac{(م ب/أ - م ب/ب - م أ/أ - م أ/ب)}{ر \times ك}$$

$$190,18 = \frac{(182,67 - 139,06)}{10}$$

ويمكن من خلال ذلك الوصول إلى قرار بشأن أفضل جمع لمستوى

للعامل (أ) مع مستوى للعامل (ب). وتحليل التباين لهذه النواتج يتبن الآتي:

مقارنة المعالجات	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة التباين
أ من خلال أو مع ب	1352,10	ر - 1 = 1	1352,10	56,93
أ مع ب	178,59	ر - 1 = 1	178,59	7,52
ب مع أ	92,48	ل - 1 = 1	92,48	3,87
ب مع أ	190,18	ل - 1 = 1	190,18	8,01
البواقي	379,92	16	23,75	

ولكن $٨,٥٣ = ١٠٠\%$ ، $٤,٤٩ = ٥٠,١٦\%$ ف

مما يمكن معه القول بأن أعلى المعالجات أثرا هي للعامل (أ) عند المستوى الأدنى للعامل (ب) ، أما أثر (أ) مع المستوى الأعلى لـ (ب) أو للعامل (ب) مع المستوى الأعلى للعامل (أ) فهو أقل معنوية. أما بالنسبة للعامل (ب) عند المستوى الأدنى للعامل (أ) فهو غير معنوي.

وجدير بالإشارة إلى أن مجموع المربعات للعامل عند مستويات العامل الآخر = مجموع المربعات للأثر البسيط أو الأساسي للعامل مضافا إليه مجموع المربعات الذي ينشأ من تفاعل هذا العامل مع العامل الآخر فمثلا:

$$١٥٣٠,٦٩ = ١٧٨,٥٩ + ١٣٥٢,١٠ = \text{م.م. (أ/ب)} + \text{م.م. (ب)}$$

$$١٥٣٠,٦٩ = ٢٧٣,٩٥ + ١٢٥٦,٧٥ = \text{م.م. (أ)} + \text{م.م. (ب)}$$

وبالمثل فإن:

$$٢٨٢,٦٦ = ١٩٠,١٨ + ٩٢,٤٨٠ = \text{م.م. (ب/أ)} + \text{م.م. (أ)}$$

$$٢٨٢,٦٦ = ٢٧٣,٩٥ + ٨,٧١ = \text{م.م. (ب)} + \text{م.م. (أ)}$$

تمارين

(١) البيانات التالية تبين عدد الوحدات المعيبة في كل ١٠٠ وحدة أنتجت في خمسة أيام متتالية في أربعة مواقع إنتاجية مختلفة:

مواقع الإنتاج	الأيام				
	(١)	(٢)	(٣)	(٤)	(٥)
أ	٧	٨	٨	١٠	١٢
ب	٥	٥	٦	٦	٨
ج	٦	٧	٨	٩	١٠
د	٥	٧	٧	٨	٨

حلل نتائج هذه التجربة باستخدام أسلوب تحليل التباين عند مستوى

المعنوية ٥% . حدد النموذج والفروض التي تختبرها ثم أوجد الحد الأدنى

للفروق المعنوية لكل من الأيام ومواقع الإنتاج عند $\alpha = ٥\%$.

(٢) لدراسة أثر أسلوب الإعلان وطريقة التعليب على الطلب على سلعة ما فقد

سجلت هذه النتائج من تجربة عشوائية على السلعة في ستة منافذ للتوزيع

لأحد المتاجر:

أسلوب الإعلان	طريقة التعليب		
	أ	ب	ج
صحافة	٣٠	٣٧	٣٢
تلفزيون	٣٤	٣٨	٤٢

تحقق من معنوية أثر أسلوب الإعلان وطريقة التعليب على الطلب على

السلعة باستخدام تحليل التباين وذلك عند مستوى المعنوية ٥% .

كيف يقارن هذا النموذج بالنموذج كامل العشوائية من حيث الكفاءة ؟.

(٣) وبفرض أن تجربة السؤال السابق قد كررت في عدة أيام متتالية فكانت

النتائج كالآتي:

طريقة التعليب			أسلوب الإعلان
أ	ب	ج	
٣٠	٣٧	٣٢	صحافة
٣١	٣٥	٣٠	
٢٨	٣٩	٢٧	
٣٤	٣٨	٤٢	تلفزيون
٣٨	٣٧	٤١	
٣٦	٤٠	٤٦	

عين النموذج الخاص بهذه التجربة، ثم حل نتائجها مستخدماً تحليل

التباين. بماذا توصي في ضوء نتائج التحليل؟

(٤) لتقدير ومقارنة ارتفاع الكفاءة لبرنامج تدريبي بحسب طول فترة التدريب،

فقد طبق البرنامج التدريبي على أربعة مجموعات متساوية العدد من

المتدربين في ثلاثة أشهر متتالية، وسجل الوقت المنفق في إنتاج سلعة ما

بالساعة فكانت النتائج كالآتي:

الشهر	المجموعة			
	(١)	(٢)	(٣)	(٤)
أبريل	١٦	١٣	١٥	٩
مايو	٣٠	٢٢	٢٩	٢٣
يونيو	٢٤	٢٠	١٨	٢٤

حلل هذه البيانات مستخدماً تحليل التباين، ثم أوجد الخطأ المعياري للفروق بين متوسطات الأشهر. قارن بين كفاءة نموذج القطاعات الكاملة العشوائية المستخدم في تصميم هذه التجربة والنموذج الكامل العشوائية. (٥) وبفرض أن القراءة الخاصة بالمجموعة الرابعة في شهر يونيو في السؤال السابق قد فقد تسجيلها، قدر هذه القيمة وأعد التحليل مرة أخرى عند نفس مستوى المعنوية.

(٦) البيانات التالية تبين الوقت (مقرباً إلى أقرب ساعة) الذي أستغرق في الإجابة على اختبارين مختلفين طبقاً في ستة أيام متتالية ، وقد أجرى الاختبار مرتين في كل حالة وكانت النتائج كالآتي:

الاختبار	اليوم					
	(١)	(٢)	(٣)	(٤)	(٥)	(٦)
أ	٧	٩	٨	٦	٨	٧
	٩	١٠	٨	٨	٩	٧
ب	٨	١٠	٧	١٠	٧	٨
	٩	١١	٨	١٠	٨	٩
المجموع	٣٣	٤٠	٣١	٣٤	٣٢	٣١

حلل نتائج هذه التجربة وناقش نتائج التحليل، وبماذا توصي في ضوء نتائج التحليل؟

(٧) أراد أحد معارض السيارات أن يقارن بين خمسة أنواع من السيارات حسب متوسط عدد الكيلومترات / لتر بنزين فأختار عينة عشوائية مكونة من ثلاث سيارات من كل نوع من ثلاث مدن كبيرة القاهرة / الإسكندرية / المنصورة وأجرى الاختبار على أساس لتر واحد من البنزين في كل سيارة من كل نوع في كل مدينة وكانت النتائج كالآتي:

المدينة	نوع السيارة			
	أ	ب	ج	د
القاهرة	٢٠,٣	١٩,٥	٢٢,١	١٧,٦
	١٩,٨	١٨,٦	٢٣,٠	١٨,٣
	٢١,٤	١٨,٩	٢٢,٤	١٨,٢
الإسكندرية	٢١,٦	٢٠,١	٢٠,١	١٩,٥
	٢٢,٤	١٩,٩	٢١,٠	١٩,٢
	٢١,٣	٢٠,٥	١٩,٨	٢٠,٣
المنصورة	١٩,٨	١٩,٦	٢٢,٣	١٩,٤
	١٨,٦	١٨,٣	٢٢,٠	١٨,٥
	٢١,٠	١٩,٨	٢١,٦	١٩,١

حلل نتائج هذه الدراسة موضحاً النموذج وفروضه.

(٨) استخدم نموذج المربع اللاتيني 4×4 لمقارنة أربعة طرق مختلفة لإنتاج سلعة ماء، وسجلت عدد الوحدات المنتجة وكانت النتائج كالآتي:

د	١٢	أ	١١	ج	١٣	ب	١٢
ب	١٤	ج	١٣	أ	١٢	د	١٥
أ	١٠	ب	٩	د	١٠	ج	١١
ج	١٢	د	١٣	ب	١١	أ	١٢

بين أن لاختلاف الطرق ليس له أثر معنوي على إنتاج. ثم قارن بين الكفاءة النسبية لهذا النموذج ونموذج القطاعات الكاملة العشوائية في ضوء نتائج التجربة وبماذا توصي لأن؟

(٩) وبفرض أن المفردة الستة تقع في الصف الرابع العمود الثاني قد سقطت تسجيلها، قنر هذه المفردة ثم أعد التحليل ثم أوجد الخطأ المعياري للفرق بين متوسط الطريقة (د) وأى طريقة أخرى.

(١٠) الآتي بيانات مربع لاتيني ٦×٦ لمقارنة ستة أنواع من البذور:

ب	و	د	أ	هـ	جـ	مجموع الصفوف
٢٢٠	٩٨	١٤٩	٩٢	٢٨٢	١٦٩	١٠١٠
أ	هـ	ب	جـ	و	د	
٧٤	٢٣٨	١٥٨	٢٢٨	٤٨	١٨٨	٩٣٤
د	جـ	و	هـ	ب	أ	
١١٨	٢٧٩	١١٨	٢٧٨	١٧٦	٦٥	١٠٣٤
هـ	ب	أ	د	جـ	و	
٢٩٥	٢٢٢	٥٤	١٠٤	٢١٣	١٦٣	١٠٥١
جـ	د	هـ	و	أ	ب	
١٨٧	٩٠	٢٤٢	٩٦	٦٦	١٢٢	٨٠٣
و	أ	جـ	ب	د	هـ	
٩٠	١٢٤	١٩٥	١٠٩	٧٩	٢١١	٨٠٨
مجموع الأعمدة	٩٨٤	١٠٥١	٩١٦	٩٠٧	٨٦٤	٥٦٤٠

حلل نتائج هذه التجربة. أي أنواع البذور أكبر أثرا على إنتاجية المحصول ثم قارن بين كفاءة هذه التجربة وكفاءتها لو أنها صممت على أساس قطاعات عشوائية كاملة.

(١١) أجريت تجربة لاختبار تباين أثر أربعة أنواع من السماد أ ، ب ، جـ ، د وذلك بترتيب النبات الذي يسد في شكل مربع لاتيني ٤×٤ أي أن الصفوف والأعمدة كانت تمثل اتجاهات مختلفة في حقل التجربة وكانت نتائج التجربة :

أ	ب	جـ	د
١٧	١٣	١٩	١٦
ب	جـ	د	أ
١٣	٢٠	١٤	١٨
جـ	د	أ	ب
٢١	١٦	١٥	١٢
د	أ	ب	جـ
١٤	١٦	١٤	١٨

حلل هذه البيانات للتحقق من أن متوسط الغلة متباين باختلاف نوع السماد وذلك عند مستوى المعنوية ٥% . ثم قارن بين كفاءة النموذج مقارنا بنموذج القطاعات العشوائية الكاملة.

(١٢) يوضح الجدول التالي التحليل الجزئي للتباين لتجربة عاملية ذات عاملين أ ب:

المصدر	م.م	د.ح	م.م.م
العامل أ	...	٣	٠,٧٥
العامل ب	٠,٩٥	١	...
التفاعل (أ ب)	٠,٣٠
الخطأ
كلى	٦,٥	٢٣	

والمطلوب:

أولاً : تحديد عدد مستويات كل من العاملين وكذلك عدد وحدات التجربة عند كل مستوى لعامل مع مستويات العامل الآخر.

ثانياً : ماذا يعنى " تفاعل عامل مع آخر " وكيف يمكن أن يستفاد منه عند تقييم نتيجة التجربة ؟

ثالثاً: استكمل جدول تحليل التباين السابق ثم اختبر معنوية الأثر المتوسط للمعالجات عند مستوى المعنوية $\alpha = 0.01$.

رابعاً : اختبر معنوية أثر كلاً من العاملين وكذلك تفاعلها عند ذات مستوى المعنوية ١% . وأى من الاختبارات السابقة (فى ثالثاً أو رابعاً) تحققها التحليل العاملى ؟

الفصل الثالث

الانحدار الخطي البسيط Simple linear Regression

محتويات الفصل :

- (١) مقدمة .
 - (٢) طرق الحصول على خط الانحدار .
 - (٣) خط الانحدار المستقيم .
 - (٤) العلاقة بين ميل خط الانحدار ومعامل الارتباط .
 - (٥) الخطأ المعياري لتقدير معادلة خط الانحدار .
 - (٦) العلاقة بين الخطأ المعياري لمعادلة خط الانحدار ومعامل الارتباط .
 - (٧) معامل التحديد ومعامل الارتباط .
 - (٨) معادلة خط الانحدار من بيانات مبوبة .
 - (٩) الاستدلال الاحصائي عن معالم خط الانحدار .
 - (١٠) تحليل التباين وتحليل الانحدار .
 - (١١) استخدام نموذج الانحدار في عملية التنبؤ بفترة ثقة .
- تمارين

الفصل الثالث

الانحدار الخطي البسيط

Simple linear Regression

(1) مقدمة :

تناولنا من قبل موضوع الارتباط سواء أكان بين ظاهرتين فقط (ارتباط بسيط) أو بين عدة ظواهر (ارتباط متعدد) وذلك بغرض دراسة شكل وطبيعة ودرجة العلاقة بين هذه الظواهر . لكن في كثير من الحالات نكون أكثر اهتماماً بشكل العلاقة التي تربط بين الظاهرتين من حيث كونها علاقة خطية (على شكل خط مستقيم) أو علاقة غير خطية (على شكل منحنى) وذلك بغرض التنبؤ بقيمة إحدى الظاهرتين ، إذا علمت قيمة الظاهرة الأخرى والبحث في شكل العلاقة بين الظواهر هو محور موضوع الانحدار .

والارتباط يختلف عن الانحدار ، فالارتباط هو دراسة التوزيع المشترك لمتغيرين أو أكثر كل يأخذ قيمة مختلفة دون تدخل من الباحث أي أن أنواع القيم المتناظرة هي قيم عشوائية ، أما الانحدار فيفترض وجود متغير واحد تابع dependent وإن القراءات المسجلة عنه عشوائية ويتأثر هذا المتغير بواحد أو أكثر من المتغيرات المستقلة (independent) حيث يتدخل الباحث بتثبيت قيم المتغير (أو المتغيرات) المستقلة عند مستويات معينة - أو كما يقال أحياناً أنها تقاس دون خطأ - ويشاهد ويسجل القيم التي يأخذها المتغير التابع عند المستويات المحددة للمتغير (أو المتغيرات) المستقلة ، ومن أنواع القيم المتناظرة في الحالة الأولى ، يمكن تقدير معامل الارتباط الذي يستدل منه على اتجاه ودرجة العلاقة بين المتغيرات موضوع الدراسة ، وفي الحالة الثانية يمكن تعيين الدالة التي تحدد العلاقة بين المتغير التابع والمتغيرات المستقلة وتستخدم هذه

الدالة فى التنبؤ بالقيم المستقبلية للمتغير التابع بدلالة قيم جديدة للمتغيرات المستقلة .

يتضح مما سبق أنه عند قياس علاقة الارتباط بين الظواهر او المتغيرات ، لا نهتم بالبحث فى أى المتغيرات تابع وأيها مستقل أو أيهما السبب وأيهما النتيجة ، بينما هذه السفرقة لها أهمية خاصة عند دراسة العلاقة الانحدارية بين المتغيرات - كما سنرى فيما بعد - وعلى ذلك يمكن القول بأن الارتباط هو وسيلة إحصائية تلخص العلاقة بين الظاهرتين (أو أكثر) فى صورة رقم (قيمته تتراوح بين -1، +1) ، أما الانحدار فهو وسيلة إحصائية تلخص العلاقة بين الظاهرتين أو أكثر فى صورة معادلة رياضية (قد تكون معادلة خط مستقيم او معادلة منحنى) بغرض تقدير قيمة المتغير التابع إذا علمت قيمة المتغير (أو المتغيرات) المستقل .

ويعتبر سير فرانسيس جالتون Sir Frances Galton أول من تعرض لموضوع الانحدار فى نهاية القرن التاسع عشر (١٨٨٥) ، فقد استخدم كلمة Regression لأول مرة عندما كان يدرس العلاقة بين أطوال مجموعة من الآباء والأبناء عددهم حوالى ١٠٠٠ وقد كشفت دراسته عن نتائج هامة ملخصها ما يلي :

- ١- أن الآباء طوال القامة لهم أبناء طوال القامة ، والآباء قصار القامة لهم أبناء قصار القامة.
- ٢- متوسط الطول للأبناء من مجموعة آباء طوال القامة أقل من أطوال آبائهم ، ومتوسط الطول للأبناء من مجموعة آباء قصار القامة أكبر من أطوال آبائهم .

وقد استخدم جالتون اصطلاح Regression (وترجمتها فى القواميس هى الارتداد أو الرجوع للخلف) قاصداً به الملاحظات السابقة . وقد ظل هذا

الاصطلاح مستخدماً حتى وقت قريب بواسطة أغلبية من الإحصائيين ، ولكن قاصدين بها الخط المرسوم لمجموعة من النقاط يمثل الاتجاه العام لها بخلاف المعنى الذي كان يقصده جالتون من استخدام هذا الاصطلاح . ولذلك نجد ان بعض الإحصائيين فى السنوات الأخيرة تفضل استخدام اصطلاح خط التقدير بدلاً من خط الانحدار . Estimating line instead of Regression line.

(٣) : طرق الحصول على خط الانحدار :

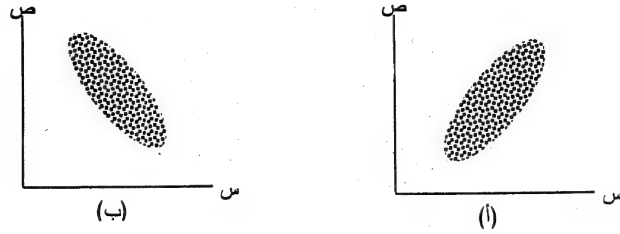
هناك طريقتين أساسيتين يمكن استخدام إحداهما للوصول إلى خط الانحدار :

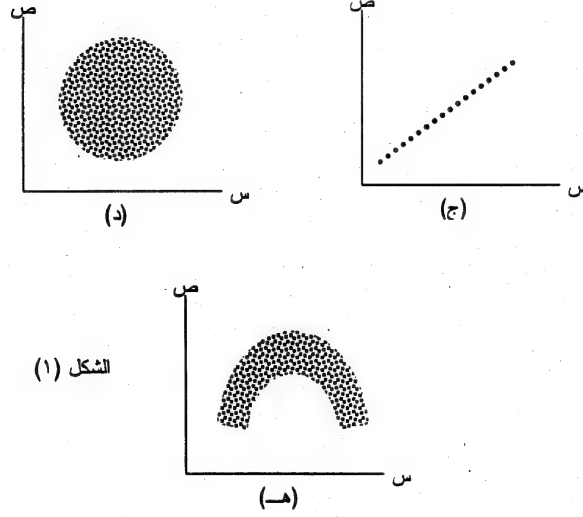
١- طريقة الشكل الانتشارى (الطريقة البيانية) .

٢- طريقة المربعات الصغرى .

(١-٣) طريقة الشكل الانتشارى : Scatter Diagram Method

من الممكن اكتشاف شكل العلاقة الفعلية التى تربط بين المتغيرين (س ، ص) بواسطة الشكل الانتشارى ، وذلك بتمثيل المتغير المستقل (س) على المحور الأفقى والمتغير التابع (ص) على المحور الرأسى . وعند رصد أزواج المشاهدات الفعلية لكل من المتغيرين (س ، ص) على لوحة بيانية يمكن أن نحصل على شكل واحد لهذه البيانات من بين عدة أشكال مختلفة موضحه بالأشكال رقم أ ، ب ، ج ، د ، هـ فى شكل (١) .





ولإيضاح مدلول المتغير المستقل والمتغير التابع نقدم بعض الأمثلة التالية :

العلاقة بين كمية الإنتاج من القطن وكمية السماد المستخدم ، هي علاقة بين متغير تابع ومتغير مستقل ، فكمية الإنتاج تتحدد تبعاً لكمية السماد المستخدم ، وبمعنى آخر بتغيير كمية السماد زيادة أو نقص يتبع ذلك تغير في كمية الإنتاج زيادة أو نقص ، أي أن التغير في كمية الإنتاج يعتمد ويتوقف على التغير الذي يحدث في كمية السماد ، لذا يقال أن السماد متغير مستقل لأن الباحث هو الذي يتحكم في اختيار مستويات أو كميات السماد المستخدمة ، أما كمية الإنتاج فهو متغير تابع لا يمكن أن يتدخل الباحث في تحديد كميته . كذلك العلاقة بين الدخل وحجم الإنفاق هي علاقة بين متغير مستقل ومتغير تابع ، فحجم الإنفاق يعتمد ويتوقف على حجم الدخل ، أيضاً العلاقة بين معدلات وفيات

الأطفال الرضع ومستوى معيشة الأسرة ، هي علاقة بين متغير تابع (معدلات الوفاة بين الأطفال الرضع) ومتغير مستقل (مستوى المعيشة).

وكما بينا من قبل نجد أنه إذا وقعت جميع النقاط على استقامة واحدة ، أي كونت خط مستقيم ، فهذا دليل على وجود ارتباط تام بين المتغيرين (سواء أكان طردي يصعد إلى أعلى أو سلبي يهبط إلى أسفل) أو قد تقع جميع النقاط على خط منحنى ، مما يعني وجود ارتباط تام بينهما . لكن التمثيل الفعلي للبيانات المستمدة من متغيرات أو ظواهر اقتصادية وتجارية لا تحقق مثل هذا الارتباط التام ، فالارتباط التام في مثل هذه الظواهر هو حالة نادرة ، وعلى ذلك من المتوقع أن نجد أن التمثيل الفعلي لآزواج القراءات (س ، ص) ينتشر ويتوزع حول خط مستقيم أو حول منحنى ، مما يعني وجود حالة من الارتباط غير التام ، وكلما اقتربت النقاط من الخط المستقيم أو من المنحنى كلما ازداد الارتباط بين الظاهرتين . والخط الذي تنتشر حوله هذه النقاط سواء كان خط مستقيم أو منحنى يسمى بخط الانحدار Regression Line وهذا الخط يمكن التعبير عنه بمعادلة رياضية سواء بمعادلة خط مستقيم على الصورة :

$$ص = أ + ب س$$
 أو بمعادلة منحنى على الصورة $ص = أ + ب س + ج س^2$ ولما كان هذا الخط ذو أهمية وفائدة كبيرة ، لأنه يستخدم في التنبؤ بقيمة المتغير التابع إذا علمت قيمة المتغير المستقل ، كما أنه يستخدم أيضا في تحديد نسبة تأثير المتغير المستقل على المتغير التابع ، فإن هذا الخط يمكن أن نصل إليه بأحدى طريقتين : طريقة التمهيد باليد أو طريقة المربعات الصغرى . فإذا استخدمنا طريقة التمهيد باليد وهي طريقة تعتمد على المهارة الشخصية للباحث وعلى خبرته في رسم الخطوط البيانية وعلى فهمه لطبيعة البيانات محل الدراسة ، فإنه يجب مراعاة النقاط التالية لكي نوفق في رسم أفضل خط انحدار سواء أكان خط مستقيم أو منحنى :

١- يجب أن يكون هذا الخط قريب جدا من جميع النقاط .

- ٢- يجب أن يقع على جانبي هذا الخط عدد متساوي من النقط بقدر المستطاع .
٣- يجب أن تكون النقط الواقعة على جانبي هذا الخط على مسافات متساوية
منه ، بمعنى أن الانحرافات الموجبة (للنقط التي فوق الخط) تتعادل مع
الانحرافات السالبة (للنقط التي تحت الخط) .

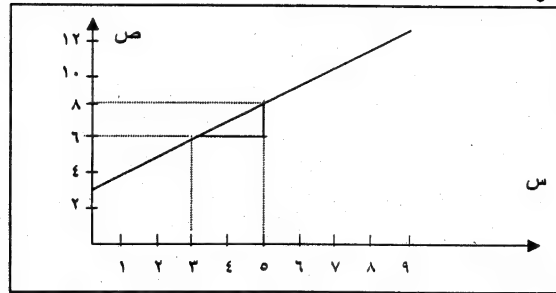
مثال (١) :

أعطيت البيانات التالية عن المتغيرين (س ، ص) حيث س : متغير مستقل ،
ص : متغير تابع

س	٢	٣	٥	٦	٨	٩
ص	٦	٥	٧	٨	١٢	١١

استخدام طريقة الشكل الانتشاري في استطلاع شكل العلاقة بين المتغيرين ، ثم
وفق أفضل خط تراه يمثل هذه البيانات .

الحل :



شكل (٢)

من الواضح أن الشكل الانتشاري للنقط عند رصدها بيانياً وفق المعايير
السابقة ، يعبر عن وجود علاقة خطية شبه مستقيمة ، يمكن التعبير عنها بخط
مستقيم على الصورة :

ص = ١ + ب س ، حيث (أ) عبارة عن الجزء المقطوع من المحور الرأسي وهو يساوي ٣ تقريبا ، أما (ب) فهي عبارة عن ميل خط الانحدار أي ظل الزاوية التي يصنعها هذا الخط مع الاتجاه الموجب لمحور السينات أو هو مقدار التغير الذي يحدث في (ص) عندما تتغير (س) بوحدة واحدة ، أي :

$$ب = \frac{ص٢ - ص١}{س٢ - س١} = \frac{٦ - ٨}{٣ - ٥} = ١ \text{ تقريبا}$$

وعلى ذلك تصبح معادلة خط الانحدار المستقيم ص = ١ + ب س على

الصورة :

$$ص = ٣ + س$$

ويمكن استخدام المعادلة السابقة في التنبؤ بقيمة المتغير التابع (ص)

عند قيمة معلومة للمتغير المستقل (س) ، فمثلا إذا كانت س = ١٠ فإن قيمة

$$(ص) \text{ المقدرة } = ١٣ = ١٠ \times ١ + ٣$$

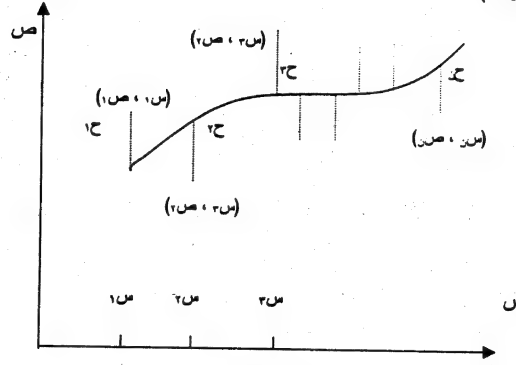
يتضح مما سبق أن التمهيد باليد طريقة سهلة وسريعة ، لكن بتعدد الأشخاص تتعدد الخطوط البيانية لنفس البيانات ، لأن كل واحد يرى أن الخط الذي رسمه للبيانات هو أفضل الخطوط من وجهة نظره ، وبالتالي ينتج عن تعدد الخطوط المستقيمة تعدد القيم المتنبأ بها للمتغير التابع (ص) عند قيمة معينة معلومة للمتغير المستقل (س) ، ذلك لأن كل خط ستختلف فيه قيمة (أ ، ب) عن الخطوط الأخرى ، وتعدد القيم التقديرية هذه دليل على عدم الدقة في توفيق معادلة تمثل هذه البيانات .

(٣-٣) طريقة المربعات الصغرى Least Square Method

لكي نتلافى أثر التقدير الشخصي عند رسم خط الانحدار ، نلجأ إلى طريقة أخرى تحكمها قواعد رياضية ثابتة لا يختلف تفسيرها أو اتباعها من شخص لآخر . لكن دعنا في البداية نوضح المقصود باصطلاح أفضل توفيق

للخط البياني Best Fitting Line (سواء خط مستقيم أو خط منحنى) .
لنفرض أنه عند التمثيل البياني لأنواع القراءات (س ، ص) حصلنا على الشكل
رقم (٣) وفيه تم رسم خط بياني يتوسط النقط الفعلية بقدر الامكان .

يلاحظ أنه عند احدى قيم المتغير المستقل س ولتكن s_1 ، أن القيمة
الفعلية المناظرة لها وهي v_1 تبعد عن الخط البياني المرسوم بالمسافة ح ،
كذلك نجد أنه عند القيمة s_2 تبعد القيمة الفعلية المناظرة لها وهي v_2 عن
الخط البياني بالمسافة ح ، وهكذا لباقي النقط . أي أن هذه الفروق أو الانحرافات
(وتسمى بالخطأ العشوائي كما سيتضح فيما بعد) عبارة عن الفرق بين القيمة
الفعلية (ص) والقيمة المقابلة لها على الخط البياني ولتكن ص (أي القيمة التي
تقرأ من على الخط البياني وتسمى بالقيمة النظرية أو المتوقعة أو بالقيمة
التقديرية)



شكل (٣)

وهذه الانحرافات بعضها موجب (إذا كانت القيمة الفعلية (ص) تقع
أعلى الخط البياني) وبعضها سالب (إذا كانت القيمة الفعلية (ص) تقع أسفل

الخط البياني (وبعضها صفر) إذا كانت القيمة الفعلية (ص) تقع على الخط البياني) وأفضل خط بياني يوفق في التعبير عن هذه البيانات ويمثلها بدقة كافية هو الخط الذي يجعل مجموع مربعات الانحرافات السابقة أقل ما يمكن أي :

$$ح^2_1 + ح^2_2 + ح^2_3 + + ح^2_n \text{ أقل ما يمكن}$$

أي مجـ ح² ر أقل ما يمكن ... (١)

وعلى ذلك إذا كانت هناك مجموعة من الخطوط البيانية ، يمكن رسمها للمتغيرين (س،ص) وكانت إحدى هذه الخطوط تحقق الشرط السابق (مجموع مربعات انحرافات القيم الفعلية عن القيم التقديرية أقل ما يمكن) فإن هذا الخط يعتبر أفضل هذه الخطوط Best Line ويسمى بخط المربعات الصغرى . ولكي نصل إلى الخط المربعات الصغرى سواء أكان خط مستقيم أو خط منحنى ، فإن طريقة المربعات الصغرى تحدد لنا الأسلوب الرياضي الذي يتبع في تحديد ثوابت كل معادلة ، وسوف نتناول هذا الأسلوب مرة عندما تكون هناك علاقة خط مستقيم بين الظاهرتين ، ومرة أخرى عندما تكون هناك علاقة منحنى من الدرجة الثانية بين الظاهرتين ، وأخيراً عندما تكون هناك علاقة خط انحدار متعدد والذي يربط بين أكثر من متغيرين .

تعود طريقة المربعات الصغرى إلى عالم الرياضيات الألماني كارل فريدريك جاوس . وفي ظل فروض معينة ، تتمتع تقديرات طريقة المربعات الصغرى بعدة خصائص احصائية هامة ، جعلتها من أدق وأفضل الطرق التي يمكن أن تستخدم في تحليل الانحدار . ولتوضيح هذه الطريقة ، نستعرض أولاً الفروض التي وضعها جاوس . بفرض أنه لدينا متغير مستقل (س) ومتغير تابع (ص) يرتبطان بعلاقة خطية على الصورة : $ص = \gamma + \beta س$. وحيث أنه من غير المتوقع - من الناحية العملية - أن تقع جميع النقاط تماماً على خط

مستقيم واحد ، فإن تلك العلاقة الخطية يجب أن تعدل كي تضم حد إضافي يسمى بالحد العشوائي (ψ) ، أي أن:

$$ص = \gamma + \beta س + \psi$$

ونموذج الانحدار بهذه الصورة يعتمد على عدة فروض ، بعضها يختص بالمتغير العشوائي (ψ) والبعض الآخر يختص بالعلاقة بين المتغير العشوائي (ψ) والمتغير المستقل ($س$) .

فروض نموذج خط الانحدار:

- ١- قيمة المتغير العشوائي (ψ) قد تكون موجبة أو سالبة أو مساوية للصفر ، لكن متوسط أو توقع هذه القيم يساوي الصفر ، أي : توقع (ψ) = صفر
- ٢- الحدود العشوائية (أو الأخطاء العشوائية) غير مرتبطة ببعضها البعض ، أي أن تغاير الحدود العشوائية يساوي الصفر ، وهذا الفرض (سوف نتناوله فيما بعد) يعني من الناحية الفنية عدم وجود ارتباط ذاتي بين الحدود العشوائية (أو اللواقح) أي No autocorrelation .
- ٣- تباين (ψ) مقدار ثابت عند جميع قيم المتغير المستقل ($س$) ، أي : $\sigma^2(\psi) = ثابت$. وهذا الفرض (سوف نتناوله فيما بعد) يعني من الناحية الفنية وجود تجانس للتباين Homoscedasticity
- ٤- قيم الأخطاء العشوائية (ψ) مستقلة (غير مرتبطة) عن قيم المتغير المستقل ($س$) ، أي:
- ٥- تغاير (ψ ، $س$) = صفر
- ٥- المتغير المستقل ($س$) مقياس بدون أخطاء ، أما المتغير التابع فقد يحتوي أولاً يحتوي على أخطاء في القياس .
- ٦- التوزيع الاحتمالي للمتغير (ψ) هو التوزيع الطبيعي ، توقعه وتباينه على الصورة :

توقع (ψ) = صفر ، $\psi^2 6$ = ثابت . ويعبر عن ذلك في الصورة المختصرة:

$\psi \sim م$ (صفر ، $\psi^2 6$) ، حيث تشير م إلى التوزيع الطبيعي المعتاد أو المعتدل.

٧- من نموذج الانحدار البسيط : $\gamma + \beta س + \psi$ = ، نجد أن المتغير التابع ص يصبح دالة خطية في المتغير العشوائي ψ ، أي أن التوزيع الاحتمالي للمتغير (ص) يتوقف ويعتمد على التوزيع الاحتمالي للمتغير (ψ) وعلى ذلك نجد أن (ص) تتبع توزيع طبيعي معتاد بتوقع وتباين على الصورة :

توقع (ص) = $\gamma + \beta س$

تباين (ص) = تباين (ψ) = $\psi^2 6$

ويعبر عن ذلك في الصورة المختصرة: $\psi \sim م$ ($\gamma + \beta س$) ، $\psi^2 6$

لاحظ أن المتغير المستقل (س) هو متغير مثبت عند مستويات معينة ، أي أنه متغير غير عشوائي أي متغير غير احتمالي ، بمعنى أنه عند سحب العديد من العينات ، فإن قيم (س) تكون ثابتة لا تتغير من عينة لأخرى ، كذلك γ مقدار ثابت .

ونظراً لأن $\beta, \gamma, \psi^2 6$ هي القيم الحقيقية (أي مؤشرات) للعلاقة بين ص ، س في مجتمع الدراسة ، وهي ثوابت عادة تكون مجهولة القيمة ، لذا فإننا نلجأ إلى تقديرها من بيانات عينة عشوائية تسحب من مجتمع الدراسة . بفرض أنه سحبنا عينة عشوائية حجمها (ن) مفردة ، وسجل لكل مفردة في تلك العينة قيم الظاهرتين (س ، ص) ، بذلك تصبح العلاقة بين (ص ، س) في عينة الدراسة على الصورة :

ص = أ + ب س + خ

حيث : (خ) يناظر (ψ) وهو يمثل الخطأ العشوائي ، أي الخطأ في قيمة (ص) ، أي الفرق بين قيمة (ص) الحقيقية أي المشاهدة والمسجلة عملياً وقيمتها المتوقعة أي الواقعة على معادلة خط الانحدار . وسوف نتناول كيفية تقدير معاملات خط الانحدار (أ ، ب) مرة عندما يكون (ص) متغير تابع أي معادلة خط انحدار ص على س ومرة أخرى عندما يكون (س) هو المتغير التابع أي معادلة خط انحدار س على ص .

(٣) خط الانحدار المستقيم Simple Linear Regression

(١-٣) معادلة خط الانحدار ص على س (ص / س) :

إذا كانت العلاقة التي تربط بين المتغيرين (س ، ص) توجي بأنها تأخذ شكل خط مستقيم - كما في شكل (٢) - معادلته على الصورة :
ص = أ + ب س + خ (٢)

حيث ص : المتغير التابع س : المتغير المستقل

أ : مقدار ثابت ، ويمثل الجزء المقطوع من المحور الرأسي وهو عبارة عن قيمة عندما تكون قيمة س = صفر .

ب : ميل أو معامل خط الانحدار أو مقدار التغير الذي يحدث في المتغير التابع (ص) عندما تتغير قيمة المتغير المستقل (س) بوحدة واحدة .

خ : الخطأ في قيمة (ص) أي اختلاف قيمة (ص) الحقيقية عن قيمتها التقديرية والواقعة على خط الانحدار .

نعلم أن الخط المستقيم يتحدد تماماً بعد معرفة قيم الثوابت (أ ، ب) وأن هذه الثوابت أو المجاهيل يمكن تحديدها بعدة طرق مختلفة ، لكن المهم هو اختيار الطريقة التي تحقق أفضل خط مستقيم يمثل انواج القراءات (س ، ص)

وفي الواقع فإن طريقة المربعات الصغرى تحدد أفضل قيم للنوابت (أ، ب)، فمن شأن هذه القيم والتي تحدد بطريقة المربعات الصغرى أن تجعل مجموع مربعات انحرافات القيم الفعلية (ص) عن القيم المتوقعة أو التقديرية (ص) أقل ما يمكن، أي: مجـ (ص - ص)² أقل ما يمكن، حيث ص هي القيم النظرية أو التقديرية أو المتوقعة أي التي تقع على خط الانحدار، أي أنها تحققه، أي أن ص = أ ب س

∴ المطلوب الآن هو تحديد قيم الثوابت أ ، ب بشرط أن :

مَجَّ حٌ = مَجَّ (ص - ص) أَقْلٌ مَا يُمْكِنُ ، أَيُّ أَنْ :

مج ح^٢ = مج (ص - أ - ب س) ^٢ نهاية صغرى .

وبإجراء عمليات التفاضل الجزئي مرة بالنسبة إلى (أ) ومرة أخرى بالنسبة إلى (ب) ومساواة الناتج بالصفر نجد أن :

$$6 \text{ مج } ح^2 = \frac{2 \text{ مج } (ص - ا - ب س) (1 -)}{16} = \text{ضر}$$

∴ ۲ - (مج - ص - ن - أ - ب - مج - س) = صفر

مج-ص-ن-أ-ب-مج-س-صفر، (-۲ = صفر)

... مجـ ص = ن أ + ب مجـ س ... (۲)

$$6 \text{ مج ح} = \frac{2 \text{ مج (ص - ا - ب س)} (س -)}{6 \text{ ب}} = \text{صفر}$$

∴ - ۲ مجس (ص - ا - ب س) = صفر

مجس ص - أ مجس - ب مجس ۲ = صفر، (۲ - = صفر)

∴ مجس ص = أ مجس + ب مجس ۲ ... (۴)

وتسمى المعادلات (٣) ، (٤) بالمعادلات المعنّاة أو الطبيعية

Normal Equations ويحل المعادلتين (٣،٤) معا بطريقة الحذف المتتالي أو

التعويض أو بالمحددات ، أو بالمصفوفات نصل إلى قيم الثوابت (أ ، ب) حيث أن : مج ص ، مج س ، مج س ص ، مج س² هي قيم معلومة .
من ناحية أخرى يمكن أن نصل من المعادلتين (٣ ، ٤) إلى علاقات رياضية للثوابت (أ ، ب) إذا استخدمنا طريقة المحددات مثلا على النحو التالي:

$$\begin{aligned} \text{مج ص} &= \text{ن} \cdot \text{أ} + \text{ب} \cdot \text{مج س} \\ \text{مج س ص} &= \text{أ} \cdot \text{مج س} + \text{ب} \cdot \text{مج س}^2 \end{aligned}$$

$$\therefore \text{أ} = \frac{\begin{vmatrix} \text{مج ص} & \text{مج س} \\ \text{مج س ص} & \text{مج س}^2 \end{vmatrix}}{\begin{vmatrix} \text{ن} & \text{مج س} \\ \text{مج س} & \text{مج س}^2 \end{vmatrix}} = \frac{\text{مج ص} \times \text{مج س}^2 - \text{مج س} \times \text{مج س ص}}{\text{ن} \cdot \text{مج س}^2 - (\text{مج س})^2} \quad (٥) \dots$$

$$\text{ب} = \frac{\begin{vmatrix} \text{ن} & \text{مج ص} \\ \text{مج س} & \text{مج س ص} \end{vmatrix}}{\begin{vmatrix} \text{ن} & \text{مج س} \\ \text{مج س} & \text{مج س}^2 \end{vmatrix}} = \frac{\text{ن} \cdot \text{مج س ص} - \text{مج س} \times \text{مج ص}}{\text{ن} \cdot \text{مج س}^2 - (\text{مج س})^2} \quad (٦) \dots$$

$$\text{أو} \quad \text{ب} = \frac{\text{مج س ص} - \text{مج س} \times \text{مج ص} / \text{ن}}{\text{مج س}^2 - (\text{مج س})^2 / \text{ن}} \quad (٧) \dots$$

وقد تمكن كل من جاوس وماركوف من اثبات أن تقديرات طريقة المربعات الصغرى (أ، ب) هي تقديرات خطية ، وغير متحيزة ، ولها أصغر تباين . وهذه الخصائص الثلاث هي من أهم الخصائص الإحصائية التي يتمتع بها أي تقدير إحصائي .

طريقة أخرى للوصول إلى المعادلات الطبيعية :

رأينا أن المعادلات الطبيعية (٣) ، (٤) هي الأساس في الوصول إلى الثوابت (أ ، ب) وقد حصلنا على هذه المعادلات ، عن طريق التفاضل الجزئي للكمية مجـ ح^٢ . لكن يمكن الوصول إلى هذه المعادلات بطريقة أخرى مبسطة يسهل تذكرها إذا تتبعنا الخطوات التالية :

$$\text{ص} - \text{أ} + \text{ب} \text{ س} \quad (\text{المعادلة الأصلية})$$

الخطوة الأولى :

اجمع طرفي المعادلة الأصلية السابقة (ن) من المرات ، وهي حجم العينة ، أي عدد انواع المشاهدات المسجلة عن الظاهرتين (س ، ص) فنحصل على أول معادلة طبيعية وهي:

$$\text{مجـ ص} - \text{ن} \text{ أ} + \text{ب مجـ س} \dots\dots\dots (\text{وهي نفس المعادلة رقم ٣})$$

الخطوة الثانية :

اضرب طرفي المعادلة الأصلية في المتغير المستقل (س) لتصبح على الصورة : س ص = أ س + ب س^٢ ثم اجمع طرفي المعادلة بعد ذلك (ن) من المرات لنحصل على المعادلة التالية :

$$\text{مجـ س ص} = \text{أ مجـ س} + \text{ب مجـ س}^2 \dots\dots\dots (\text{وهي نفس المعادلة رقم ٤})$$

صور أخرى بديلة للتقديرات (أ ، ب) :

من المعادلة (٣) نلاحظ ما يلي: مجـ ص = ن أ + ب مجـ س بقسمة طرفي المعادلة على (ن)

$$\frac{\text{مجـ ص}}{\text{ن}} = \text{أ} + \text{ب} \frac{\text{مجـ س}}{\text{ن}} \quad \text{أي} \quad \text{ص} - \text{أ} + \text{ب} \text{ س}$$

$$\therefore \text{أ} = \text{ص} - \text{ب س} \quad \dots\dots\dots (٨)$$

ومن المعادلة (٤) وباستخدام طريقة انحرافات القيم الفعلية لكل من س ، ص

عن أوساطهما الحسابية $\bar{ص}$ ، $\bar{س}$ نجد أن :

$$\text{مج س ص} = \text{أ مج س} + \text{ب مج س}^2$$

$$\text{مج (س - } \bar{س} \text{) (ص - } \bar{ص} \text{) = أ مج (س - } \bar{س} \text{) + ب مج (س - } \bar{س} \text{) + ب مج}$$

$$\text{مج (س - } \bar{س} \text{) (ص - } \bar{ص} \text{) = ٢ (ص - } \bar{ص} \text{) + ب مج (س - } \bar{س} \text{) + ٢ (س - } \bar{س} \text{)}$$

[مج (س - $\bar{س}$) (ص - $\bar{ص}$) = صفر لأن مجموع انحرافات القيم عن وسطها الحسابي = صفر]

$$\text{ب} = \frac{\text{مج (س - } \bar{س} \text{) (ص - } \bar{ص} \text{)}}{\text{مج (س - } \bar{س} \text{)}^2} \quad (٩) \dots$$

مثال (٣)

مستخدماً بيانات مثال (١) ، أوجد معادلة خط انحدار ص على س بطريقة المربعات الصغرى.

الحل :

معادلة خط انحدار ص على س (ص / س) هي :

$$\text{ص} = \text{أ} + \text{ب س}$$

ويمكن تحديد الثوابت (أ ، ب) بطريقة المربعات الصغرى إما بحل

مجموعة المعادلات الطبيعية معاً أو باستخدام العلاقات الرياضية وكلاهما يعطي

نفس النتيجة ، وسوف نعرض للحلين بغرض التوضيح :

جدول (١)

س	ص	س ص	س ^٢	ص ^٢
٢	٦	١٢	٤	٣٦
٣	٥	١٥	٩	٢٥
٥	٧	٣٥	٢٥	٤٩
٦	٨	٤٨	٣٦	٦٤
٨	١٢	٩٦	٦٤	١٤٤
٩	١١	٩٩	٨١	١٢١
٣٣	٤٩	٣٠٥	٢١٩	٤٣٩

تحديد (أ، ب) باستخدام المعادلات الطبيعية :

$$\text{ص} = \text{أ} + \text{ب س}$$

$$\text{مجم ص} = \text{ن أ} + \text{ب مجم س}$$

$$\text{مجم س ص} = \text{أ مجم س} + \text{ب مجم س}^2$$

بالتعويض بالمجاميع الموضحة بجدول (١)

$$٤٩ = ٣٣ + ١٦ \text{ ب}$$

$$٣٠٥ = ٢١٩ + ١٣٣ \text{ ب}$$

بضرب المعادلة الأولى في (٣٣) وبضرب المعادلة الثانية في (٦) ثم الطرح نجد أن :

$$١٦١٧ = ١٩٨ + ١٠٨٩ \text{ ب}$$

$$١٨٣٠ = ١٩٨ + ١٣١٤ \text{ ب}$$

$$٢١٣ - ٢٢٥ = \text{صفر} - ٢٢٥ \text{ ب}$$

$$٢١٣ - ٢٢٥ = \text{ب}$$

وبالتعويض عن قيمة (ب) في أي معادلة ولتكن الأولى نصل إلى قيمة (أ)

$$٤٩ = ٣٣ \times ٠,٩٥ + ١٦$$

$$٣١,٣٥ + ١٦ =$$

$$٤٩ - ٣١,٣٥ = ١٦ - ١٧,٦٥ \quad ١٦ = ١٧,٦٥ - ١ \quad ١٦ - ١٧,٦٥ = -١,٦٥$$

معادلة خط الانحدار ص = أ + ب س بعد تحديد الثوابت أ، ب تصبح

على الصورة :

$$\text{ص} = ٠,٩٥ + ٢,٩٤ \text{ س}$$

وبلاحظ أن المعادلة التي توصلنا إليها بالطريقة البيانية من قبل قريبة

جدا من هذه المعادلة .

تحديد (أ، ب) باستخدام العلاقات الرياضية :

$$\text{ص} = \text{أ} + \text{ب س} , \text{ حيث :}$$

$$\begin{aligned} \frac{49 \times 33}{6} - 30.0 &= \frac{\text{مـ س} \times \text{مـ ص}}{\text{ن}} - \frac{\text{مـ س ص} - \text{مـ س}^2}{\text{ن}} \\ \frac{1(33)}{6} - 219 &= \frac{(\text{مـ س})^2}{\text{ن}} - \frac{\text{مـ س ص} - \text{مـ س}^2}{\text{ن}} \\ 0.90 &= \frac{30.0}{37.0} = \frac{69.0 - 30.0}{181.0 - 219} = \text{ب} \\ \frac{\text{مـ س}}{\text{ن}} \times 0.90 - \frac{\text{مـ ص}}{\text{ن}} &= \text{أ} - \text{ص} - \text{ب س} \\ 0.0 \times 0.90 - 8.17 &= \frac{33}{6} \times 0.90 - \frac{49}{6} = \text{ب} \\ 2.94 &= 0.23 - 8.17 = \end{aligned}$$

∴ معادلة خط الانحدار هي :

$$\text{ص} = 0.90 + 2.94 \text{ س}$$

وهي نفس المعادلة السابقة .

صور أخرى لمعاملات خط الانحدار بطريقة الانحرافات البسيطة والمعدلة :

كما ذكرنا من قبل ، يمكن تبسيط العمليات الحسابية اللازمة لتقدير

النوابت (أ، ب) وخاصة (ب) عن طريق اختيار أوساط فرضية من كل من س ،

ص ، ولتكن ٠.٠ و ٠.٢

... الطريقة المباشرة

$$\begin{aligned} \frac{\text{مـ س} \times \text{مـ ص}}{\text{ن}} - \frac{\text{مـ س ص} - \text{مـ س}^2}{\text{ن}} &= \text{ب} \\ \frac{(\text{مـ س})^2}{\text{ن}} - \frac{\text{مـ س ص} - \text{مـ س}^2}{\text{ن}} &= \text{ب} \end{aligned}$$

فإذا استخدمنا طريقة الانحرافات البسيطة نجد أن :

$$\begin{aligned} \frac{\text{مـ ح}^2 \times \text{مـ ح}^2}{\text{ن}} - \frac{\text{مـ ح}^2 \text{ ح}^2 - \text{مـ ح}^2}{\text{ن}} &= \text{ب} \\ \frac{(\text{مـ ح}^2)^2}{\text{ن}} - \frac{\text{مـ ح}^2 \text{ ح}^2 - \text{مـ ح}^2}{\text{ن}} &= \text{ب} \end{aligned}$$

حيث $\bar{C} = \bar{S} - \bar{O}$ ، $\bar{C} = \bar{V} - \bar{O}$ ، أما الأرقام (١ ، ٢) فترمز للمتغيرات (س ، ص) على التوالي . وإذا كانت الانحرافات البسيطة \bar{C}_1 ، \bar{C}_2 كل يقبل القسمة على عامل مشترك \bar{C}_1 ، \bar{C}_2 على التوالي فإن :

$$b = \frac{\bar{C}_1 \bar{C}_2 - (\bar{C}_1 \bar{C}_2 \times \bar{C}_1 \bar{C}_2 / N)}{(\bar{C}_1 \bar{C}_2 - (\bar{C}_1 \bar{C}_2 / N))}$$

$$= \frac{(\bar{C}_1 \bar{C}_2 - (\bar{C}_1 \bar{C}_2 \times \bar{C}_1 \bar{C}_2 / N))}{(\bar{C}_1 \bar{C}_2 - (\bar{C}_1 \bar{C}_2 / N))} \quad (11)$$

أما الثابت (أ) فيكفى استخدام المعادلة رقم (٨) حيث :

$$a = \bar{S} - b \bar{S} = \bar{V} - \frac{\bar{C}_1 \bar{C}_2}{N} \times \bar{C}_1 \bar{C}_2$$

مثال (٣) :

الجدول التالي يوضح كميات الإنتاج (ص) من محصول القمح الناتجة عن كميات مختلفة من السماد (س) في عينة من ١٠ أفدنة :

كمية الإنتاج: ص	٤٠	٤٤	٤٦	٤٨	٥٢	٥٨	٦٠	٦٨	٧٤	٨٠
كمية السماد: س	٦	١٠	١٢	١٤	١٦	١٨	٢٢	٢٤	٢٦	٣٢

المطلوب توفيق أفضل معادلة خط انحدار ص على س بطريقة

المربعات الصغرى ثم ارسم هذا الخط بيانياً - ما هي كمية الإنتاج المتوقعة إذا كانت كمية السماد = ٢٠ وحدة .

الحل :

بدلاً من استخدام القيم الحقيقية لكل من س ، ص في إيجاد معاملات خط الانحدار أ ، ب يمكن استخدام طريقة الانحرافات البسيطة وذلك بطرح وسط فرضي (و١) من قيم (س) وليكن $\bar{O} = ٢٠$ ويطرح وسط فرضي (و٢) من قيم (ص) وليكن $\bar{O} = ٥٠$ ، بذلك نحصل على الانحرافات \bar{C}_1 ، \bar{C}_2 حيث $\bar{C}_1 =$

س - ٢٠ ، ح - ص = ٥٠ وبالتالي بدلاً من إجراء العمليات الحسابية السابقة على (س ، ص) نجرى على بدائلها ح ، ٢٠ .

جدول (٢)

س	ص	ح	٢ ح	١ ح	٢ ح	١ ح
٦	٤٠	١٤-	١٠-	١٠-	١٩٦	١٤٠
١٠	٤٤	١٠-	٦-	٦-	١٠٠	٦٠
١٢	٤٦	٨-	٤-	٤-	٦٤	٣٢
١٤	٤٨	٦-	٢-	٢-	٣٦	٢٨
١٦	٥٢	٤-	٢+	٢+	١٦	١٦
١٨	٥٨	٢-	٨+	٨+	٤	٤
٢٢	٦٠	٢+	١٠+	١٠+	٤	٤
٢٤	٦٨	٤+	١٨+	١٨+	١٦	٣٢
٢٦	٧٤	٦+	٢٤+	٢٤+	٣٦	٥٧٦
٣٢	٨٠	١٢+	٣٠+	٣٠+	١٤٤	٩٠٠
١٨٠	٥٧٠	٢٠-	٧٠+	٧٠+	٦١٦	٢١٢٤

معادلة خط الانحدار ص / س : ص = أ + ب س حيث :

$$ب = \frac{\frac{\sum \text{مـ ح} \times \text{مـ ح}}{ن} - \frac{\sum \text{مـ ح} \times \text{مـ ح}}{ن}}{\frac{\sum (\text{مـ ح})^2}{ن} - \frac{(\sum \text{مـ ح})^2}{ن^2}}$$

$$ب = \frac{\frac{70 \times 20}{10} - \frac{140 + 816}{40 - 616}}{\frac{(20)^2}{10} - \frac{140^2}{4000}} = \frac{140 - 816}{40 - 616} = \frac{956}{576} = 1,66$$

$$أ = \text{ص} - ب س = \frac{\sum \text{مـ ص}}{ن} - 1,66 \times \frac{\sum \text{مـ س}}{ن}$$

$$27,12 = 29,88 - 0,57 = 18 \times 1,66 - 0,57 = \frac{180}{10} \times 1,66 - \frac{0,57}{10} =$$

∴ معادلة خط الانحدار هي :

$$\text{ص} = 1,66 + 27,12 \text{ س}$$

تقدير كمية الإنتاج عندما تكون كمية السماد = 20 وحدة هي :

$$\text{ص} = 1,66 + 27,12 \times 20 = 33,20 + 27,12 = 60,32 \text{ وحدة}$$

ولرسم معادلة خط الانحدار السابقة ، نحتاج إلى معرفة الاحداثيات السينية والصادية لنقطتين فقط . فمثلاً :

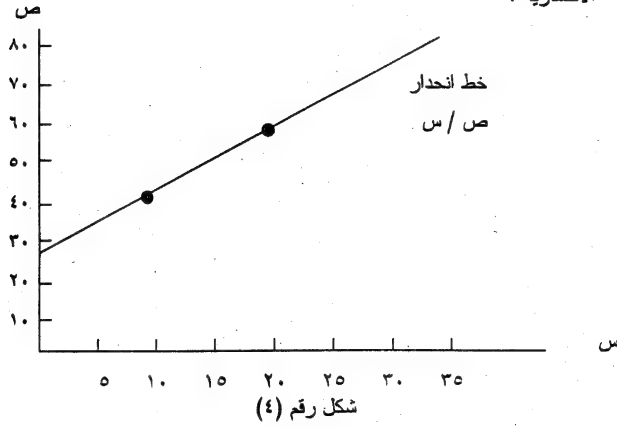
$$\text{عندما تكون س} = 10 \text{ فإن ص} = 10 \times 1,66 + 27,12 = 43,72$$

$$\text{وعندما تكون س} = 20 \text{ فإن ص} = 20 \times 1,66 + 27,12 = 60,32$$

∴ النقطتان هما : (10 ، 43,72) ، (20 ، 60,32)

فى الشكل رقم (٤) تم رصد انواع المشاهدات (النقط الانتشارية) وتم

رسم خط الانحدار التقديرى ومن الواضح أن هذا الخط يقترب جداً من النقط الانتشارية .



ملاحظات :

- ١- يلاحظ أن الأوساط الحسابية لكل من s ، \bar{y} أعداد صحيحة ($s = ١٨$ ، $\bar{y} = ٥٧$) ومن ثم يكون من الأفضل اختيار الأوساط الفرضية $s = ١٠$ ، $\bar{y} = ٥٠$ مساوية للأوساط الحسابية $s = ١٨$ ، $\bar{y} = ٥٧$ ، وفي هذا تخفيض كبير في حجم العمليات الحسابية ، حيث نستخدم المعادلة رقم (٩) لحساب قيمة (ب) ، أما (أ) فلا خلاف في القانون الرياضي المستخدم لها وهو دائماً العلاقة رقم (٨).
- ٢- يلاحظ أيضاً أن الانحرافات البسيطة s ، \bar{y} في هذا المثال كل منها يمكن تبسيطه أكثر من ذلك بالقسمة على عامل مشترك $s = ٢$ ، $\bar{y} = ٢٩$ حيث نجد أن $s = ٢$ ، $\bar{y} = ٢٩$ (ليس من الضروري أن تكون $s = ٢$ ، $\bar{y} = ٢٩$) وفي هذا أيضاً تخفيض كبير في حجم العمليات الحسابية ، حيث نستخدم المعادلة رقم (١١) لحساب قيمة (ب) . ويترك للقارئ إعادة الحل مرة أخرى مراعيًا الملاحظات (١ ، ٢) ليتأكد أنه سيصل حتماً لنفس معادلة خط الانحدار السابقة .

(٣-٣) معادلة خط انحدار s على y (s/y) :

هناك بعض المتغيرات تتبدل فيها طبيعة المتغير المستقل والمتغير التابع، بمعنى قد يصبح المتغير المستقل هو المتغير التابع أو العكس . فمثلاً العلاقة بين الطول والوزن قد يكون أحدهما هو التابع والآخر هو المستقل . إذا افترضنا أن هناك علاقة خطية من الدرجة الأولى تربط ما بين المتغير المستقل (y) ، وهو المتغير الذي تتحدد قيمة مسبقاً دون تدخل من الباحث ، والمتغير التابع (s) وهو المتغير الذي تتحدد قيمة تبعاً للتغير الذي يحدث في المتغير المستقل (y) فإن معادلة خط الانحدار s على y (s/y) تكون على الصورة :

$$s = a + b y \quad \dots (١٢)$$

حيث : s : متغير مستقل ، y : متغير تابع .

جـ : مقدار ثابت ، وهو عبارة عن الجزء المقطوع من المحور الأفقي (محور السينات) أو هو عبارة عن قيمة (س) عندما تكون ص = صفر .

د : ميل أو معامل خط انحدار س على ص وهو عبارة عن مقدار التغير الذي يحدث في المتغير التابع (س) إذا تغير المتغير المستقل (ص) بوحدة واحدة.

وأفضل تقدير للثوابت (جـ ، د) هي تلك التي نحصل عليها بطريقة المربعات الصغرى ، وهي الطريقة التي تجعل مجموع مربعات انحرافات القيم الفعلية للمتغير التابع (س) عن القيم المتوقعة (س) والواقعة على خط الانحدار أقل ما يمكن . ويلاحظ ان الانحرافات هنا هي انحرافات أفقية ، عكس الحال في معادلة خط انحدار (ص على س) حيث كانت الانحرافات هناك رأسية . وبتتبع طريقة المربعات الصغرى ، سنجد أننا نهدف إلى جعل الكمية التالية :

$$مج - ح^2 = مج - (س - س^2) \text{ أقل ما يمكن}$$

وبما أن القيم التقديرية س تقع على خط الانحدار فهي إذن تحققها ومن ثم نجد أنه مطلوب جعل مج - (س - ج - د ص) أقل ما يمكن بإجراء عمليات التفاضل الجزئي للكمية السابقة مرة بالنسبة إلى (جـ) ومرة أخرى بالنسبة إلى (د) ومساواة النتائج في كل مرة بالصفر ، نصل إلى المعادلات الطبيعية التالية :

$$مج - س = ن - ج + د مج - ص \dots (١٣)$$

$$مج - س = ج - مج - ص + د مج - ص^2 \dots (١٤)$$

وبحل المعادلتين (١٣ ، ١٤) معاً بأي طريقة جبرية نصل إلى قيمة الثوابت (جـ ، د) أو يمكن استخدام العلاقات الرياضية التالية :

$$\dots \text{س} = \text{ج} + \text{د ص}.$$

$$\dots (١٥) \quad \text{د} = \frac{\text{مج س ص} - \frac{\text{مج س} \times \text{مج ص}}{\text{ن}}}{\text{مج ص}^2 - \frac{(\text{مج ص})^2}{\text{ن}}} \quad \text{أو :}$$

$$\dots (١٦) \quad \text{د} = \frac{\text{مج (س - ص)} (\text{س} - \text{ص})}{\text{مج (ص - ص)}^2}$$

$$\dots (١٧) \quad \text{ج} = \text{س} - \text{د ص}$$

$$\text{حيث : } \text{س} = \frac{\text{مج س}}{\text{ن}}, \quad \text{ص} = \frac{\text{مج ص}}{\text{ن}}$$

وإذا استخدمنا طريقة الانحرافات البسيطة أو المعدلة بغرض تبسيط العمليات الحسابية نجد أن ميل خط الانحدار (د) يأخذ الصور التالية :

$$\dots (١٨) \quad \text{د} = \frac{\text{مج ح}^2 - \frac{\text{مج ح} \times \text{مج ح}^2}{\text{ن}}}{\text{مج ح}^2 - \frac{(\text{مج ح})^2}{\text{ن}}}$$

أو :

$$\dots (١٩) \quad \text{د} = \frac{\text{ث}^2}{\text{مج ح}^2 - \frac{(\text{مج ح})^2}{\text{ن}}} \times \frac{\text{مج ح}^2 - \frac{\text{مج ح} \times \text{مج ح}^2}{\text{ن}}}{\text{مج ح}^2 - \frac{(\text{مج ح})^2}{\text{ن}}}$$

حيث (ن) في جميع العلاقات السابقة هي حجم العينة ، أي عدد أزواج القراءات المتناظرة بين (س ، ص) .

ملاحظات :

- ١- يلاحظ أن بسط ميل خط الانحدار (ب أو د) في جميع المعادلات السابقة يساوي بسط معامل الارتباط الخطي البسيط لبيرسون .
- ٢- يلاحظ أن هناك تماثل وتطابق كامل بين بسطي ميل خط الانحدار (ب ، د).
- (٣-٣) العلاقة بين معادلتَي خط الانحدار :**

بينما من قبل أن معادلتَي خط الانحدار (ص على س) ، (س على ص) هما على الصورة التالية:

$$\begin{aligned} \text{ص} &= \text{أ} + \text{ب س} & \dots (\text{ص} / \text{س}) \\ \text{س} &= \text{ج} + \text{د ص} & \dots (\text{س} / \text{ص}) \end{aligned}$$

فإذا أردنا تقدير قيمة ص (متغير تابع) عند قيمة محددة للمتغير المستقل س ، فإننا نستخدم معادلة خط انحدار (ص/س) وإذا أردنا تقدير قيمة س (كمتغير تابع) عند قيمة محددة للمتغير المستقل ص ، فإننا نستخدم معادلة خط انحدار (س/ص) .

والمعادلتين ص = أ + ب س ، س = ج + د ص ، يختلفان عن بعضهما اختلافاً كاملاً، ففي المعادلة الأولى يتم تقدير الثوابت (أ ، ب) بطريقة المربعات الصغرى ، بحيث تجعل الانحرافات الرأسية بين القيم الفعلية للمتغير التابع (ص) والقيم التقديرية (ص) الواقعة على خط الانحدار أقل ما يمكن ، بينما في المعادلة الثانية يتم تقدير الثوابت (ج ، د) بطريقة المربعات الصغرى أيضاً ، بحيث تجعل الانحرافات الأفقية للمتغير التابع (س) والقيم التقديرية (س) الواقعة على خط الانحدار أقل ما يمكن .

وهناك علاقة رياضية تربط ما بين معادلتَي خط الانحدار يمكن إيضاحها على النحو التالي :

$$\text{ب} = \frac{\text{مج} - (\text{س} - \text{ص})(\text{ص} - \text{ص})}{\text{مج} - (\text{س} - \text{س})(\text{س} - \text{س})} = \frac{\text{مج} - \text{ص} - \text{مج} \text{س} \times \text{مج} \text{ص} / \text{ن}}{\text{مج} - (\text{مج} \text{س} - \text{س} / \text{ن})}$$

$$\dots د = \frac{\text{مـجـ} (س - \bar{س}) (ص - \bar{ص})}{\text{مـجـ} (ص - \bar{ص})} = \frac{\text{مـجـ} س ص - \text{مـجـ} س \bar{ص} - \text{مـجـ} \bar{س} \bar{ص} + \text{مـجـ} \bar{س} \bar{ص}}{\text{مـجـ} (ص - \bar{ص})} = \frac{\text{مـجـ} س ص - \text{مـجـ} \bar{س} \bar{ص}}{\text{مـجـ} (ص - \bar{ص})}$$

$$\dots ب \times د = \frac{[\text{مـجـ} (س - \bar{س}) (ص - \bar{ص})]^2}{\text{مـجـ} (ص - \bar{ص}) \times \text{مـجـ} (ص - \bar{ص})} = \frac{[\text{مـجـ} (س - \bar{س}) (ص - \bar{ص})]^2}{\text{مـجـ} (ص - \bar{ص})^2}$$

$$= \frac{\text{مـجـ} س ص - \text{مـجـ} \bar{س} \bar{ص}}{\text{مـجـ} (ص - \bar{ص})} = \frac{(\text{مـجـ} س - \bar{س}) (\text{مـجـ} ص - \bar{ص})}{\text{مـجـ} (ص - \bar{ص})} = \frac{(\text{مـجـ} س - \bar{س}) (\text{مـجـ} ص - \bar{ص})}{\text{مـجـ} (ص - \bar{ص})}$$

ونائج حاصل الضرب السابق ما هو إلا مربع معامل الارتباط الخطى البسيط لبيرسون أما الجذر التربيعي لحاصل ضرب (ب × د) فيعطى معامل الارتباط الخطى (ر) . نخلص من هذا إلى أن :

$$\dots (٢٠) \quad \boxed{ر = \sqrt{ب \times د}}$$

أي أن الجذر التربيعي لحاصل ضرب معاملتي خط الانحدار (ب ، د) ، يعطى معامل الارتباط (ر) بين الظاهرتين (س ، ص) وهذه علاقة إضافية لحساب معامل الارتباط العادي .

ملاحظات :

- ١- يلاحظ عند رسم خط الانحدار (ص/س) ، (س/ص) في شكل واحد نجد ، أنهما يتقاطعان عند النقطة (س ، ص) ، بمعنى أن الإحداثي الأفقي لنقطة التقاطع هو الوسط الحسابي للظاهرة (س) والإحداثي الرأسى لنقطة التقاطع هو الوسط الحسابي للظاهرة (ص) ويرجع ذلك إلى النقطة (س،ص) تحقق كلتا المعادلتين أى تقع على خطى الانحدار .
- ٢- يلاحظ أنه إذا كان هناك ارتباط طردي موجب بين الظاهرتين (س ، ص) ، فإن خطى الانحدار (ص/س) ، (س/ص) يميل كل منهما على المحور الأفقي ميلاً موجباً أي أن كل منهما يصنع زاوية حادة مع محور السينات (المحور الأفقي) .

- ٣- إذا كان الارتباط بين الظاهرتين ارتباط عكسي سالب ، فإن خطي الانحدار يصنع كل منهما زاوية منفرجة مع المحور الأفقي .
- ٤- أما إذا كان هناك ارتباط تام موجب أو سالب بين الظاهرتين ، فإن خطي الانحدار (ص/س) (س/ص) ينطبقان على بعضهما البعض .
- ٥- وإذا كان الارتباط بين الظاهرتين منعدم (ر = صفر) ، فإن خطي الانحدار يتعامدان مع بعضهما ويصنعا بينهما زاوية قائمة = ٩٠° .

مثال (٤)

الجدول التالي يوضح أوزان عينة مكونة من ١٢ أب (س) وكبر الأبناء (ص) .

(س)	٦٥	٦٣	٦٧	٦٤	٦٨	٦٢	٧٠	٦٦	٦٨	٦٧	٦٩	٧١
(ص)	٦٨	٦٦	٦٨	٦٥	٦٩	٦٦	٦٨	٦٥	٦١	٦٧	٦٨	٧٠

المطلوب :

- ١- أوجد معادلة خط انحدار ص على س .
- ٢- أوجد معادلة خط انحدار س على ص .
- ٣- من معادلتى خط الانحدار ، استنتج قيمة معامل الارتباط بين طول الأب وطول الابن .
- ٤- اعرض بيانياً معادلتى خط الانحدار بجانب النقاط الانتشارية للقراءات الفعلية.

الحل :

يمكن استنتاج معادلتى خط الانحدار إما باستخدام القراءات الأصلية ، وفى هذا مجهود حسابي ضخم أو باستخدام انحرافات القراءات الأصلية عن الأوساط الحسابية لكل من س ، ص وفى هذا مجهود حسابي ضخم أيضاً لأن الأوساط الحسابية هنا أعداد كسرية . ولكن من الأفضل استخدام انحرافات القراءات الأصلية عن أوساط فرضية و١ ، و٢ لكل من س ، ص فإذا فرض أن و١ = ٦٠ و٢ = ٦٠ فإننا يمكن أن نكون الجدول التالي :

١- معادلة خط انحدار ص على س : ص = أ + ب س حيث :

$$ب = \frac{\text{مجم ح}^2 - \frac{\text{مجم ح} \times \text{مجم ح}^2}{ن}}{\text{مجم ح}^2 - \frac{(\text{مجم ح})^2}{ن}}$$

$$أ = \frac{٦٤٧ - \frac{٩١ \times ٨٠}{١٢}}{١٢} = \frac{٦٠٦,٦٧ - ٦٤٧}{١٢} = \frac{٤٠,٣٣}{١٢} = ٣,٣٦$$

جدول (٣)

س	ص	ح	ح ^٢	ح ^٣	ح ^٤
٦٥	٦٨	٥	٨	٢٥	٤٠
٦٣	٦٦	٣	٦	٢٧	١٨
٦٧	٦٨	٧	٨	٤٩	٥٦
٦٤	٦٥	٤	٥	١٦	٢٥
٦٨	٦٩	٨	٩	٦٤	٥١٢
٦٢	٦٦	٢	٦	٨	١٢
٧٠	٦٨	١٠	٨	١٠٠	٨٠
٦٦	٦٥	٦	٥	٢١٦	٣٠
٦٨	٧١	٨	١١	٦٤	٨٨
٦٧	٦٧	٧	٧	٤٩	٤٩
٦٩	٦٨	٩	٨	٨١	٧٢
٧١	٧٠	١١	١٠	١٢١	١١٠
٨٠٠	٨١١	٨٠	٩١	٦١٨	٧٢٩

أ = ص - ب س

$$35,83 = 31,75 - 67,58 = \frac{800}{12} \times 0,476 - \frac{811}{12} =$$

∴ معادلة خط الانحدار ص/س هي : $ص = 0,476 + 35,83$

٢- معادلة خط انحدار س على ص : $س = ج + د ص$ حيث :

$$د = \frac{\frac{\sum (ج \times ص)}{ن} - \frac{\sum ج}{ن} \times \frac{\sum ص}{ن}}{\frac{\sum (ج^2)}{ن} - \frac{(\sum ج)^2}{ن^2}}$$

$$\frac{40,33}{38,92} = \frac{40,33}{690,8 - 729} = \frac{\frac{91 \times 80}{12} - 647}{\frac{729}{12} - \frac{7(91)^2}{12}}$$

ج = س - د ص

$$3,35 = 70,02 - 66,67 = \frac{811}{12} \times 1,036 - \frac{800}{12} =$$

∴ معادلة خط انحدار س / ص هي :

(٣) استنتاج قيمة معامل الارتباط : $س = -1,036 + 3,35$

∴ ص = $0,476 + 35,83$ س (ص/س)

س = $-1,036 + 3,35$ ص (س/ص)

أي أن : ب = $0,476$ ، د = $-1,036$

∴ ر = $\sqrt{ب \times د}$

$$0,7 = \sqrt{0,493} = \sqrt{1,036 \times 0,476}$$

ملحوظة :

بما أن قيمة معامل الارتباط يجب أن تكون أقل من الواحد الصحيح ، فلا بد وأن يكون أحد معاملي خط الانحدار (ب أو د) أقل من الواحد الصحيح .

(٤) معادلتى خط الانحدار بيانيا :

لرسم أي معادلة انحدارية ، تحتاج إلى معرفة إحداثيات أي نقطتين اختياريين .

• لرسم معادلة خط انحدار ص / س :

$$\text{ص} = ٠,٤٧٦ + ٣٥,٨٣ \text{ س}$$

$$\text{إذا كانت س} = ٦٠ \text{ فإن ص} = ٠,٤٧٦ + ٣٥,٨٣ = ٦٤,٣٩$$

$$\text{وإذا كانت س} = ٧٠ \text{ فإن ص} = ٠,٤٧٦ + ٣٥,٨٣ = ٦٩,١٥$$

∴ النقطتان هما : (٦٤,٣٩ ، ٦٠) ، (٦٩,١٥ ، ٧٠) .

وهذه المعادلة موضحة بالخط رقم (١) في شكل (٥)

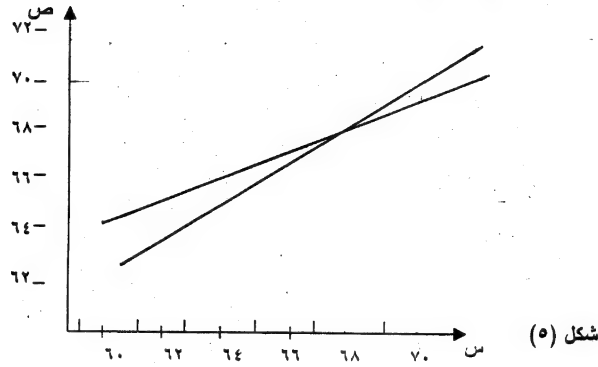
• ولرسم معادلة خط انحدار س / ص :

$$\text{س} = - ٣,٣٥ + ١,٠٣٦ \text{ ص}$$

$$\text{إذا كانت ص} = ٦٥ \text{ فإن س} = - ٣,٣٥ + ١,٠٣٦ \times ٦٥ = ٦٣,٩٩$$

$$\text{وإذا كانت ص} = ٧٠ \text{ فإن س} = - ٣,٣٥ + ١,٠٣٦ \times ٧٠ = ٦٩,١٧$$

∴ النقطتان هما : (٦٣,٩٩ ، ٦٥) ، (٦٩,١٧ ، ٧٠) .



وهذه المعادلة موضحة بالخط رقم (٢) في شكل (٥)
 ويلاحظ أن إحداثيات نقطة التقاطع هي (٦٦,٧ ، ٦٧,٦) وهذه الإحداثيات ما
 هي إلا الأوساط الحسابية لكل من \bar{S} ، \bar{V} على التوالي أي $\bar{S} = ٦٦,٧$ ،
 $\bar{V} = ٦٧,٦$

(٤) العلاقة بين ميل خط الانحدار ومعامل الارتباط:

١- معادلة خط انحدار \bar{S} على \bar{V} (\bar{S} / \bar{V}) :

$$\bar{S} = A + B \bar{V}$$

$$B = \frac{\text{مـجـ}(\bar{S} - \bar{S}) (\bar{V} - \bar{V})}{\text{مـجـ}(\bar{S} - \bar{S})^2} \quad (\text{ميل خط الانحدار})$$

.. معامل الارتباط الخطي البسيط لبيرسون بين الظاهرتين (\bar{S} ، \bar{V})

على الصورة :

$$R = \frac{\text{مـجـ}(\bar{S} - \bar{S}) (\bar{V} - \bar{V})}{\sqrt{\text{مـجـ}(\bar{S} - \bar{S})^2 \times \text{مـجـ}(\bar{V} - \bar{V})^2}}$$

$$\text{بضرب العلاقة الأخيرة في الكمية } \frac{\sqrt{\text{مـجـ}(\bar{S} - \bar{S})^2}}{\sqrt{\text{مـجـ}(\bar{S} - \bar{S})^2}} \text{ نصل إلى:}$$

$$R = \frac{\text{مـجـ}(\bar{S} - \bar{S}) (\bar{V} - \bar{V})}{\sqrt{\text{مـجـ}(\bar{S} - \bar{S})^2} \times \sqrt{\text{مـجـ}(\bar{V} - \bar{V})^2}}$$

(٢١)...

$$R = \frac{\text{مـجـ} \bar{S} \times \bar{V}}{\text{مـجـ} \bar{S} \times \bar{V}}$$

حيث \bar{S} ، \bar{V} عبارة عن الانحراف المعياري لكل من \bar{S} ، \bar{V} على التوالي
 . والعلاقة الأخيرة يمكن إعادة صياغتها على النحو التالي :

(٢٢)...

$$\frac{ب}{ع} = ر \times \frac{ع}{ع}$$

العلاقات الأخيرة (٢١) ، (٢٢) تعنى أنه يمكن إيجاد قيمة (ر) إذا علمت قيمة (ب) والعكس صحيح . وعلى ذلك يمكن إعادة صياغة معادلة خط انحدار ص على س بدلالة معامل الارتباط (ر) على النحو التالي :

$$ص = أ + ب س$$

... أ = ص - ب س حيث س ، ص هما الأوساط الحسابية لكل من الظاهرتين (س ، ص) .

$$\therefore ص = (ص - ب س) + ب س$$

(٢٣)...

$$\therefore (ص - ص) = (ب - ب س) (س - س)$$

$$وبما أن \frac{ب}{ع} = ر \times \frac{ع}{ع}$$

(٢٤)...

$$\therefore (ص - ص) = (ب - ب س) \frac{ع}{ع} (س - س)$$

أو على الصورة التالية :

(٢٥)...

$$\frac{ص - ص}{ع} = \frac{ب - ب س}{ع} ر$$

(٢) معادلة خط انحدار ص على ص (س / ص) :

$$ص = ج + د ص$$

باتباع نفس المنهج السابق يمكن أن نصل إلى النتائج التالية :

(٢٦)...

$$ر = د \times \frac{ع}{ع}$$

(٢٧)...

$$د = ر \times \frac{ع}{ع}$$

كما أنه يمكن إعادة صياغة معادلة خط انحدار س / ص السابقة بدلالة

معلومية معامل الارتباط (ر) على النحو التالي :

$$س = ج + د ص$$

$$س = (س - د ص) + د ص$$

(٢٨)...

$$(س - د ص) = د (ص - ص)$$

$$د = ر \times \frac{ع}{ع}$$

(٢٩)...

$$(س - د ص) = ر \times \frac{ع}{ع} (ص - ص)$$

(٣٠)...

$$س - د ص = ر \times \frac{ص - ص}{ع}$$

أي أنه بمعلومية كل من ر ، س ، ص ، ع ، يمكن استنتاج معادلة خط انحدار س / ص بالمعادلة (٢٤) أو معادلة خط انحدار س / ص بالمعادلة (٢٩)

مثال (٥) :

أجرى باحث دراسة تحليلية باستخدام العينة على ظاهرتين (س ، ص) وتوصل إلى المعلومات التالية عن معادلتى خط الانحدار :

$$٨ س - ١٠ ص + ٦٦ = صفر ، ٤ س - ١٨ ص = ٢١٤ ، تبين الظاهرة (س) = ٩$$

المطلوب :

- ١- إيجاد قيمة الأوساط الحسابية \bar{S} ، $\bar{ص}$.
- ٢- معامل الارتباط بين الظاهرتين ($س$ ، $ص$).
- ٣- الانحراف المعياري للظاهرة ($ص$).

الحل :

(١) إيجاد قيمة الأوساط الحسابية \bar{S} ، $\bar{ص}$:

.. معادلات الانحدار معطاة على الصورة التالية :

$$٨ \text{ س} - ١٠ \text{ ص} = -٦٦ ، \quad ٤٠ \text{ س} - ١٨ \text{ ص} = ٢١٤$$

وحيث أن خطي الانحدار يمرا دائما بالنقطة ($\bar{س}$ ، $\bar{ص}$) ، بمعنى أن النقطة ($\bar{س}$ ، $\bar{ص}$) تحقق كل معادلة .

$$\therefore ٨ \bar{س} - ١٠ \bar{ص} = -٦٦ ، \quad ٤٠ \bar{س} - ١٨ \bar{ص} = ٢١٤$$

وبحل هذه المعادلات أنيا عن طريق ضرب الأولى في (٥) ثم الطرح من الثانية ، نجد أن :

$$٤٠ \bar{س} - ٥٠ \bar{ص} = -٣٣٠$$

$$\frac{٤٠ \bar{س} - ١٨ \bar{ص}}{٣٢ \bar{ص}} = \frac{٢١٤}{٥٤٤} \quad \therefore \bar{ص} = \frac{٥٤٤}{٣٢} = ١٧$$

وبالتعويض في أي معادلة نحصل على $\bar{س}$

$$٨ \bar{س} - ١٠ \bar{ص} = -٦٦ \quad \therefore ٨ \bar{س} - ١٧ \times ١٠ = -٦٦$$

$$٨ \bar{س} = -٦٦ + ١٧٠ = ١٠٤ ، \quad \bar{س} = ١٠٤ \div ٨ = ١٣$$

$$\bar{س} = ١٣ ، \quad \bar{ص} = ١٧$$

(٢) استنتاج قيمة معامل الارتباط (r) :

نفرض أن المعادلة الأولى تمثل انحدار $ص$ / $س$:

$$\therefore \bar{ص} = \frac{٦٦}{١٠} + \frac{٨}{١٠} \bar{س} ، \therefore \text{ميل خط الانحدار (ب)} = \frac{٨}{١٠}$$

ولنفرض أن المعادلة الثانية تمثل انحدار س / ص :

$$\text{س} = \frac{18}{40} + \frac{214}{40} \text{ ص} \quad \therefore \text{ميل خط الانحدار (د)} = \frac{18}{40}$$

$$\therefore \text{ر} = \sqrt{\text{ب} \times \text{د}} = \sqrt{\frac{18}{40} \times \frac{8}{10}} = 0,6$$

(٣) إيجاد الانحراف المعياري للظاهرة (ص) :

$$\therefore \text{ميل خط الانحدار ص على س} = \frac{8}{10} = \text{ب}$$

$$\therefore \text{ب} = \text{ر} \times \frac{\text{ع}}{\text{س}} \quad \text{حيث ع س} = \sqrt{9} = 3$$

$$\frac{8}{10} = \frac{\text{ع}}{3} \times 0,6 \quad \therefore \text{ع س} = \frac{8}{10} \times 3 \times \frac{10}{6} = 4$$

مثال (٦) :

بفرض أنه توفرت لديك البيانات التالية :

الظاهرة (ص)	الظاهرة (س)	
١٠٠	١٨	الوسط الحسابي
٢٠	١٤	الانحراف المعياري
		معامل الارتباط بين س ، ص = ٠,٨ +

المطلوب :

(١) القيمة المتوقعة للمتغير (ص) إذا كانت س = ٧٠

(٢) القيمة المتوقعة للمتغير (س) إذا كانت ص = ٩٠

الحل :

١- للتنبؤ بقيمة (ص) نستخدم معادلة خط انحدار ص على س . وفي ظل المعلومات المتوفرة في هذا المثال ، نستخدم إحدى صور معادلة خط انحدار ص / س التي سبق ذكرها وتتوفر فيها هذه المعلومات .

$$\therefore (ص - \bar{ص}) = r \times \frac{ع - \bar{ع}}{ع س} (س - \bar{س})$$

$$(ص - ١٠٠) = ٠,٨ \times \frac{٢٠}{١٤} (س - ١٨)$$

$$\therefore ص = ١,١٤ س - ٢٠,٥٢ + ١٠٠$$

$$\therefore ص = ١,١٤ + ٧٩,٤٨ س$$

القيمة المتوقعة لـ ص عندما تكون س = ٧٠ هي :

$$ص = ١,١٤ \times ٧٠ + ٧٩,٤٨ = ١٥٩,٢٨$$

لاحظ أن القيمة ١٥٩,٢٨ هي القيمة المتوقعة للمتغير ص عندما تكون قيمة س = ٧٠، بمعنى أنه لو كان هناك العديد من قيم س ، كل منها تأخذ القيمة ٧٠، نجد أن هناك العديد من قيم ص المختلفة ، لكن متوسط تلك القيم (أي القيمة المتوقعة) هو ١٥٩,٢٨ وهي القيمة الواقعة على خط الانحدار .

٢- للتنبؤ بقيمة (س) ، نستخدم معادلة خط انحدار س على ص . وفي ظل المعلومات المتوفرة في هذا المثال ، نستخدم الصورة التالية من صور معادلة خط انحدار س / ص

$$(س - \bar{س}) = r \times \frac{ع - \bar{ع}}{ع س} (ص - \bar{ص})$$

$$(س - ١٨) = ٠,٨ \times \frac{١٤}{٢٠} (ص - ١٠٠)$$

$$س = ٠,٥٦ ص - ١٨ + ٥٦$$

القيمة المتوقعة لـ س عندما تكون ص = ٩٠ هي

$$س = ٠,٥٦ \times ٩٠ - ١٨ + ٥٦ = ١٢,٤$$

مثال (٧):

استخدم المعلومات التالية في إيجاد معادلة خط انحدار ص على س :

الظاهرة (ص)	الظاهرة (س)	
١٥	١٥	حجم العينة .
١٨	٢٥	الوسط الحسابي .
١٢٨	١٣٦	مجموع مربعات انحرافات القيم عن وسطها الحسابي .
---	١٢٢	مجموع حاصل ضرب انحرافات (س) ، (ص) عن أوساطهما الحسابية .

الحل:

معادلة خط انحدار ص / س هي :

$$\text{ص} = \text{أ} + \text{ب س} \quad \text{حيث } \text{ب} = \frac{\text{ع}}{\text{د}}$$

$$\text{لكن } \text{ر} = \frac{\text{مجم} (\text{س} - \text{م} \text{ن}) (\text{ص} - \text{م} \text{ن})}{\text{مجم} (\text{س} - \text{م} \text{ن})^2} = \frac{\text{ع}}{\text{د}}$$

$$\text{ع} = \text{م} \text{ن} - \frac{1}{\text{ن}} (\text{م} \text{ن} - \text{م} \text{ن})^2 \quad \text{... (٣١)}$$

$$3,02 = 9,141 - \frac{1}{14} \times 128 = 9,14 - \frac{1}{14} \times 128$$

$$\text{ع} = \text{م} \text{ن} - \frac{1}{\text{ن}} (\text{م} \text{ن} - \text{م} \text{ن})^2 = 9,71 - 136 \times \frac{1}{14} = 9,71 - 9,711$$

$$3,12 = 9,711 - \frac{1}{14} \times 128$$

$$\text{ب} = \frac{\text{ع}}{\text{د}} = \frac{3,02}{3,12} = 0,968$$

$$0,896 = \frac{3,02}{3,12} \times 0,925 = 0,896$$

$$\text{أما } \bar{a} = \bar{b} \text{ س}$$

$$= 18 - 0.896 \times 20 = 18 - 17.92 = 0.08$$

∴ معادلة خط الانحدار ص / س تصبح على الصورة :

$$\text{ص} = 0.896 \text{ س} + 0.08$$

(5) الخطأ المعياري لتقدير معادلة خط الانحدار :

ذكرنا من قبل أنه من الأهداف الرئيسية لمعادلة خط الانحدار ، هو استخدامها في عملية التنبؤ بقيمة المتغير التابع عند قيمة معلومة للمتغير المستقل . وطالما أن الارتباط بين الظواهر الطبيعية غير تام بصفة عامة ، فلا بد وأن تنحرف بعض القسيم الفعلية للمتغير التابع عن القيم المتوقعة أو التقديرية أو النظرية ، وهي القيم التي نقدرها من معادلة خط الانحدار . وطريقة المربعات الصغرى - كما ذكرنا من قبل - تهدف أساسا إلى تصغير هذه الانحرافات والوصول بها إلى أقل قيمة ممكنة . لكن يكون من المهم والمفيد قياس مدى تشتت هذه الانحرافات ، أو بمعنى أدق قياس تشتت القيم الفعلية للمتغير التابع عن قيمها المتوقعة أي عن خط الانحدار ، وهذا القياس هو ما نعني به الخطأ المعياري standard error ، ويختلف مفهوم الانحراف المعياري (الذي سبق دراسته) عن الخطأ المعياري ، فالانحراف المعياري deviation standard يقيس تشتت المفردات حول وسطها الحسابي بينما الخطأ المعياري يقيس تشتت التقديرات حول خط الانحدار .

(أ) الخطأ المعياري لتقدير معادلة خط الانحدار ص / س : ص = أ + ب س

لقياس تشتت القيم الفعلية (ص) حول خط الانحدار ، نستخدم العلاقة التالية :

... (32)

$$\text{خ}^2 \text{ ص} = \frac{\text{مجم} (\text{ص} - \hat{\text{ص}})^2}{\text{ن} - 2}$$

وهذه العلاقة توضيح تباين القيم الفعلية (\hat{y}) عن القيم المتوقعة y ،
والناتجة عن معادلة خط انحدار (y / x) ، والجذر التربيعي يعطي الخطأ
المعياري للتقدير .

ومن الواضح أنه كلما كانت قيمة x من كبيرة ، كلما دل ذلك على تباعد
أو تشتت مفردات المتغير التابع عن خط الانحدار ، مما يعني في نفس الوقت
انخفاض قيمة معامل الارتباط بين الظاهرتين (r ، y) والعكس صحيح ،
وسوف نوضح فيما بعد العلاقة التي تربط بين الخطأ المعياري ومعامل
الارتباط .

جديد بالذكر أن مقام التباين في المعادلة الأخيرة (٣٢) ما هو إلا
درجات الحرية ، وهي عبارة عن حجم العينة مطروحا منها عدد المعالم التي
يتم تقديرها من العينة وهي (a ، b) أي أن درجات الحرية = $n - 2$. وقد
درج البعض ، بدافع السهولة إلى استخدام (n) فقط بدلا من ($n - 2$) ، إلا
أن هذا الاستخدام يتسم بعدم الدقة .

ولحساب الخطأ المعياري الموضح بالمعادلة الأخيرة (٣٢) تتبع
الخطوات التالية :

- ١- أوجد معادلة خط انحدار y / x وهي $y = a + bx$
- ٢- أوجد القيم المتوقعة أو النظرية (\hat{y}) باستخدام المعادلة السابقة عند
قيم x مختلفة .
- ٣- احسب الفرق بين القيم الفعلية y والمتوقعة \hat{y} ثم رتب هذا الفرق
($y - \hat{y}$)
- ٤- أوجد مجموع مربعات الفروق السابقة $\sum (y - \hat{y})^2$
- ٥- طبق علاقة الخطأ المعياري x من التالية :

$$x = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}} \quad (33) \dots$$

وفي الواقع فإننا لسنا مضطرين دائما لحساب القيم النظرية (\hat{y}) من معادلة خط الانحدار لكي نصل إلى قيمة الخطأ المعياري ، لأن هذا يشكل صعوبة بالغة في إجراء العمليات الحسابية ، ولكن يمكن أن نستعين بعلاقة أخرى أكثر بساطة لحساب الخطأ المعياري وهي :

$$s.e. = \frac{\sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}}}{\sqrt{\sum x^2}}$$

حيث $s.e.$: مج y ، \hat{y} : مج y من بيانات موجودة بالفعل وسبق استخدامها عند إيجاد معادلة خط الانحدار ، كما أن a ، b قيم تم تقديرها .

ملحوظة :

في علم الإحصاء يفرق دائما بين الانحراف المعياري والخطأ المعياري ، فالانحراف المعياري يقصد به الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي ، أما الخطأ المعياري فيقصد به الجذر التربيعي لمتوسط مجموع مربعات انحرافات التقديرات (التي نحصل عليها من العديد من العينات) عن المتوسط العام لهذه التقديرات (متوسط المجمع) . وبمعنى أكثر اختصارا ، الانحراف المعياري يقيس درجة تشتت المفردات حول وسطها الحسابي أما الخطأ المعياري فيقيس درجة تشتت التقديرات حول متوسطها العام .

(ب) الخطأ المعياري لتقدير معادلة خط انحدار $y = a + bx$: $s.e. = \frac{\sqrt{\sum (y - \hat{y})^2}}{\sqrt{\sum x^2}}$ لقياس تشتت قيم المتغير التابع (y) حول خط الانحدار ، أي بعداً عن القيم التقديرية \hat{y} ، نستخدم العلاقة التالية :

$$s.e. = \frac{\sqrt{\sum (y - \hat{y})^2}}{\sqrt{\sum x^2}}$$

حيث \hat{S} هي القيم التقديرية الناتجة عن استخدام معادلة خط انحدار S/V ، و يمكن استخدام علاقة أخرى مماثلة للعلاقة رقم (٣٤) وهي :

$$\chi^2_{S} = \frac{\text{مج } S^2 - \frac{(\text{مج } S)^2}{N}}{N - 2} \dots (36)$$

و الخطأ المعياري هو الجذر التربيعي للتباين في المعادلات (٣٥) ، (٣٦)

مثال (٨) :

استخدم بيانات المثال رقم (٣) في إيجاد الخطأ المعياري لتقديرات معادلة خط انحدار S/V

الحل :

من مثال رقم (٣) وجدنا أن معادلة خط انحدار S/V كانت على الصورة التالية :

$$S = 1.66 + 27.12 V$$

وكما ذكرنا من قبل ، فهناك صيغتان لحساب الخطأ المعياري ، الأولى بالمعادلة رقم (٣٣) و الثانية بالمعادلة (٣٤) وسوف نعرض الحلين بغرض توضيح كيفية الاستخدام.

$$(1) \quad \text{خطأ} = \sqrt{\frac{\text{مج } (S - \hat{S})^2}{N - 2}}$$

وهنا يلزم إيجاد القيم التقديرية \hat{S} باستخدام معادلة خط الانحدار ، وذلك عند جميع قيم (S) المعطاة في المثال ، و الجدول رقم (٤) يصور العمليات الحسابية اللازمة لقياس الخطأ المعياري .

جدول (٤)

س	ص	ص - ص̂ = ٢٧,١٢ + ١,٦٦ س	ص - ص̂	(ص - ص̂)²
٦	٤٠	٣٧,٠٨ = ٦ × ١,٦٦ + ٢٧,١٢	٢,٩٢ +	٨,٥٢٦٤
١٠	٤٤	٤٣,٧٢ = ١٠ × ١,٦٦ + ٢٧,١٢	٠,٢٨ +	٠,٠٧٨٤
١٢	٤٦	٤٧,٠٤ =	١,٠٤ -	١,٠٨١٦
١٤	٤٨	٥٠,٣٦ =	٢,٣٦ -	٥,٥٦٩٦
١٦	٥٢	٥٣,٦٨ =	١,٦٨ -	٢,٨٢٢٤
١٨	٥٨	٥٧,٠٠ =	١ +	١,٠٠٠٠
٢٢	٦٠	٦٣,٦٤ =	٣,٦٤ -	١٣,٢٤٩٦
٢٤	٦٨	٦٦,٩٦ =	١,٠٤ +	١,٠٨١٦
٢٦	٧٤	٧٠,٢٨ =	٣,٧٢ +	١٣,٨٣٨٤
٣٢	٨٠	٨٠,٢٤ = ٣٢ × ١,٦٦ + ٢٧,١٢	٠,٤٢ -	٠,٠٥٧٦
			صفر	٤٧,٣٠٥٦

$$\therefore \text{الخطأ المعياري لـ } \bar{y} = \frac{\frac{47,3056}{8}}{\sqrt{2-1.0}} = \frac{47,3056}{\sqrt{1.0}} = 2,43 = 0,9132 \sqrt{1.0}$$

$$2,43 = 0,9132 \sqrt{1.0}$$

$$(ب) \text{ لـ } \bar{y} = \frac{\text{مجم ص}^2 - \frac{(\text{مجم ص})^2}{\text{ن}}}{\text{مجم ص} - \text{ب}} = \frac{13,8384 - \frac{(47,3056)^2}{8}}{26 - 10} = \frac{13,8384 - 2,8384}{16} = \frac{11,0000}{16} = 0,6875$$

وهنا يلزم معرفة \bar{y} ، \bar{y} ، \bar{y} ، \bar{y} ، ولم تكن قد حسبت في جدول (٢) أما $\bar{y} = 27,12$ ، $\bar{y} = 1,66$.

جدول (٥)

س	ص	س ص	ص ^٢
٦	٤٠	٢٤٠	١٦٠٠
١٠	٤٤	٤٤٠	١٩٣٦
١٢	٤٦	٥٥٢	٢١١٦
١٤	٤٨	٦٧٢	٢٣٠٤
١٦	٥٢	٨٣٢	٢٧٠٤
١٨	٥٨	١٠٤٤	٣٣٦٤
٢٢	٦٠	١٣٢٠	٣٦٠٠
٢٤	٦٨	١٦٣٢	٤٦٢٤
٢٦	٧٤	١٩٢٤	٥٤٧٦
٣٢	٨٠	٢٥٦٠	٦٤٠٠
١٨٠	٥٧٠	١١٢١٦	٣٤١٢٤

$$\therefore \text{خط} = \frac{11216 \times 1,66 - 570 \times 27,12 - 34124}{2 - 10} = \frac{47,04}{8} = 5,88$$

$$2,42 - 5,88 = -3,46$$

وهي نفس النتيجة السابقة ولكن مع اختلاف ضئيل نتيجة للقيم التقريبية لكل من أ ، ب .

(٦) العلاقة بين الخطأ المعياري لمعادلة خط الانحدار ومعامل الارتباط:

بيننا من قبل أنه إذا كانت النقط الانتشارية قريبة من خط الانحدار ، دل ذلك على وجود علاقة ارتباط قوية بين الظاهرتين (س ، ص) والعكس صحيح ،

فكلما تباعدت النقط الانتشارية عن خط الانحدار كلما دل ذلك على ضعف علاقة الارتباط بين الظاهرتين . تشتت النقط حول خط الانحدار المستقيم عبارة عن الخطأ المعياري والذي تناولناه في البند السابق .
من تلك الملاحظة نجد أن هناك علاقة عكسية بين الخطأ المعياري ومعامل الارتباط، فكلما كان تشتت النقط حول خط الانحدار (الخطأ المعياري) قليلاً ، كلما كان معامل الارتباط كبيراً والعكس صحيح . ومن الممكن ترجمة هذه النتيجة في صورة علاقة رياضية سواء أكان خط الانحدار هو ص/س أو س/ص .

أ- إذا كانت علاقة خط الانحدار هي : ص = أ + ب س :

$$\text{خ}^2 \text{مر} = \text{ع}^2 \text{مر} (1 - r^2) \quad (37) \dots$$

أى أن

$$\text{خ} \text{مر} = \text{ع} \text{مر} \sqrt{1 - r^2} \quad (38) \dots$$

أى أنه إذا علم معامل الارتباط (ر)، يمكن استنتاج الخطأ المعياري غير وهذه صورة أخرى من صور الخطأ المعياري بجانب المعادلات رقم (٣٣، ٣٤).
ومن الممكن استنتاج قيمة معامل الارتباط بدلالة معلومية الخطأ المعياري من العلاقة (٣٨) على النحو التالى :

$$r = \sqrt{1 - \frac{\text{خ}^2 \text{مر}}{\text{ع}^2 \text{مر}}} \quad (39) \dots$$

حيث ع^٢ مر في المعادلات الأخيرة هي تباين المفردات الفعلية للمتغير التابع ص ، أى أن :

$$\text{ع}^2 \text{مر} = \frac{1}{n} (\text{ص} - \text{ص}^2) = \frac{1}{n} (\text{مجم ص}^2 - \frac{(\text{مجم ص})^2}{n}) \quad (40) \dots$$

وعلى ذلك : ع^٢ م : تباين مفردات (ص) حول متوسطها ص .
خ^٢ م : تباين مفردات (ص) حول خط الانحدار ص/س .

ب- إذا كانت معادلة خط الانحدار هي : ص = جـ + د ص :
فإن العلاقة التي تربط بين الخطأ المعياري ومعامل الارتباط تكون على الصورة:

$$\text{خ م} = \sqrt{1 - \frac{\text{ع م}}{\text{ع م}^2}} \quad (٤١) \dots$$

حيث ع م : الانحراف المعياري لمفردات المتغير التابع س عن الوسط الحسابي م .

خ م : الخطأ المعياري لمفردات المتغير التابع عن خط الانحدار أي عن القيم م .

ر^٢ : مربع معامل الارتباط ويسمى بمعامل التحديد (وسوف نتناوله تفصيلاً فيما بعد) وجذره التربيعي هو معامل الارتباط .

من ناحية أخرى يمكن استنتاج قيمة (ر) بدلالة خ م :

$$ر = \sqrt{1 - \frac{\text{خ م}^2}{\text{ع م}^2}} \quad (٤٢) \dots$$

ويلاحظ من المعادلات (٣٨ ، ٤١) أنه في حالة الارتباط التام (طردى

أو عكسي) بين الظاهرتين س ، ص أي أن : ر = +١ فإننا نجد :

$$\text{خ م} = \text{ع م} \times \text{ص م} = \text{ص م} \times \text{ص م} = \text{ص م}^2$$

$$\text{خ م} = \text{ع م} \times \text{ص م} = \text{ص م} \times \text{ص م} = \text{ص م}^2$$

أي أنه في حالة الارتباط التام بين الظاهرتين ، فإن خطأ التقدير أو

الخطأ المعياري يساوي الصفر لأن الارتباط التام يعني أن جميع النقاط تقع على

خط الانحدار (ص/س أو س/ص) وبالتالي لا يوجد تشتت لقيم المتغير التابع .

مثال (٩):

مستخدماً نتائج مثال (٨) قدر قيمة معامل الارتباط (ر) بين كمية الإنتاج (ص) وكمية السماد (س).

الحل:

من نتائج مثال (٨) توصلنا إلى أن الخطأ المعياري لمفردات المتغير التابع ص
أى $\bar{X}_ص = ٢,٤٣$ ، $\bar{X}_س = ٥,٨٨$.

$$r = \frac{\bar{X}_ص - 1}{\bar{X}_س - 1} = \dots$$

من الواضح أننا نحتاج إلى معرفة قيمة $\bar{X}_ص$:

$$\dots \bar{X}_ص = \frac{1}{n} \sum (ص - \bar{ص})^2 = \frac{1}{n} \sum (ص - \bar{ص})^2 = \frac{1}{n} \sum (ص - \bar{ص})^2$$

ومن الجدول (٥) نجد أن :

$$\bar{X}_ص = \frac{1}{9} (34124 - \frac{570}{10}) = \dots$$

$$181,06 = \frac{1634}{9} = (32490 - 34124) \frac{1}{9} = \dots$$

$$r = \frac{5,88}{181,06} - 1 = -0,968 = -0,968$$

وللتأكد من صحة هذه النتيجة نقوم بحساب معامل الارتباط من بيانات

المثال الأصلي وهو مثال (٣) وباستخدام الجدول (٢) :

$$r = \frac{\sum (ص - \bar{ص})(س - \bar{س})}{\sqrt{\sum (ص - \bar{ص})^2 \sum (س - \bar{س})^2}} = \dots$$

$$\frac{140 + 816}{(490 - 2124)(40 - 616)} = \frac{(10/70 \times 20 -) - 816}{[10/70(70) - 2124][10/70(20) - 616]} =$$

$$0,980 = \frac{906}{970,10} = \frac{906}{16934 \times 0,576} =$$

وهي نفس النتيجة التي حصلنا عليها من قبل ، ولكن مع اختلاف طفيف حوالي ٠,١ % .

مثال (١٠) :

إذا كانت معادلتى خط انحدار ص/س ، س/ص على النحو التالي :

$$\text{ص} = 10 + 0,3 \text{ س} \quad \dots \text{ (ص/س)}$$

$$\text{س} = -4 + 1,6 \text{ ص} \quad \dots \text{ (س/ص)}$$

فما هي قيمة الخطأ المعياري لتقديرات معادلة خط انحدار س على ص ، إذا علمت أن تباين مفردات الظاهرة (س) = ٤ .

الحل :

من معادلتى خط الانحدار يمكن استنتاج قيمة معامل الارتباط .

$$\therefore r^2 = b \times d = 0,3 \times 1,6 = 0,48$$

$$\therefore \text{خ ب} = \text{ع ب} = \sqrt{1 - r^2} \quad \text{حيث} \quad \text{ع ب} = \sqrt{4 - 0,48}$$

$$\therefore \text{خ ب} = 2 - 0,48 = 1,52 \quad \text{ع ب} = 2 - 0,48 = 1,52$$

(٧) معامل التحديد ومعامل الارتباط :

فى بداية الحديث عن موضوع الانحدار ، ذكرنا أن من أهداف دراسة هذا الموضوع هو معرفة درجة تأثير المتغير المستقل (س مثلاً) على المتغير التابع (ص) ، فقد يكون مهما معرفة درجة تأثير السماد (متغير مستقل) على حجم الإنتاج (متغير تابع) . نعلم أن حجم الإنتاج يتوقف ويعتمد على عدة

عوامل منها السماد ، كمية المياه ، درجة الحرارة ، درجة الرطوبة ، مهارة العامل ، نوع التربة ... إلخ . فإذا قسمنا مجموعة هذه العوامل إلى قسمين : قسم يضم السماد (وهو المتغير المستقل الذي نريد معرفة درجة تأثيره على المتغير التابع) وقسم ثانى يضم باقى العوامل الأخرى ، ويطلق عليها اسم مجموعة العوامل العشوائية ، نجد أن التغير الذي يحدث فى الإنتاج يرجع إلى : التغير الذي يحدث فى المتغير المستقل (السماد) والتغير الذي يحدث فى العوامل العشوائية .

يطلق على التغير الذي يحدث فى الإنتاج ، أى فى المتغير التابع ، بالتغير الكلى Total Variation أما التغير الذي يحدث فى المتغير المستقل ، فيطلق عليه اسم التغير المفسر Explained Variation ، بينما التغير الذي يحدث فى المتغيرات أو العوامل العشوائية يسمى بالتغير غير المفسر Unexplained Variation . أى أن التغيرات الكلية فى الإنتاج (متغير تابع) زيادة أو نقص ، تتكون من شقين : شق يمكن تفسيره على أنه ناتج عن وجود السماد (متغير مستقل) أثر فى كمية الإنتاج ، وشق آخر لا يمكن تفسيره بعامل محدد ، ولكنه بصفة عامة يرجع إلى خليط من عوامل عشوائية لا يمكن التحكم فيها (وهو افتراض فى ظل هذه التجربة وإن كان يمكن إخراج بعض المتغيرات المستقلة من هذه العوامل العشوائية مثل كمية المياه ودرجة الحرارة كمعامل مؤثرة فى الإنتاج بجانب السماد وهذا ننقلنا إلى موضوع آخر وهو الانحدار المتعدد) .

∴ التغير الكلى فى المتغير التابع = التغير المفسر (بسبب وجود المتغير المستقل) + التغير غير المفسر (عوامل عشوائية) .

جدير بالذكر أنه فى علم الإحصاء نفرق بين اصطلاحى التغير والتباين Variation & Variance فالتغير يقصد به مجموع مربعات انحرافات القيم عن وسطها الحسابي (بالرموز : التغير = مج (ص - ص̄)² ، أما التباين

فيقصد به متوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي .
[بالرموز : التباين = مج (ص - ص̄)² / (ن-١)] . في ضوء هذه التفرقة نجد أن :

١- التغير الكلي في المتغير التابع (ص) = مج (ص - ص̄)² ... (٤٣)

$$= مج ص - \frac{(مج ص)²}{ن} \dots (٤٤)$$

٢- التغير المفسر الناشئ عن وجود المتغير المستقل (س) أي الناشئ عن استخدام معادلة خط

الانحدار ص/س والتي تعطى القيم التقديرية ص̂ = مج (ص - ص̄)²

$$= ب مج (س - س̄) (ص - ص̄) \dots (٤٥)$$

$$= ب (مج س ص - \frac{مج س \times مج ص}{ن}) \dots (٤٦)$$

٣- التغير غير المفسر والذي يرجع إلى العوامل العشوائية ، وهو عبارة عن الفرق بين القيم الفعلية (ص) والقيم التقديرية ص̂ والتي حصلنا عليها من خط الانحدار = مج (ص - ص̂)²

...التغير الكلي في (ص) = التغير المفسر بسبب وجود س + التغير غير المفسر (الخطأ العشوائي)

$$مج (ص - ص̄)² = مج (ص - ص̂)² + مج (ص̂ - ص̄)²$$

يلاحظ أنه عند ثبات حجم التغير الكلي ، فإن أي نقص في حجم الخطأ العشوائي ، يعني زيادة في حجم التغير المفسر الناشئ عن استخدام علاقة خط الانحدار ، مما يعني في نفس الوقت زيادة تأثير المتغير المستقل على المتغير التابع ، ومما يعني أيضاً زيادة قيمة معامل الارتباط بين المتغيرين (س،ص) والعكس صحيح .

ويعرف معامل التحديد Coefficient of Determination على أنه نسبة للتغير المفسر إلى التغير الكلي ويرمز له بالرمز r^2 أى أن :

$$\text{معامل التحديد } r^2 = \frac{\text{التغير المفسر}}{\text{التغير الكلي}}$$

$$= \frac{\text{مج } (\hat{y} - \bar{y})^2}{\text{مج } (y - \bar{y})^2} \quad \dots (٤٧)$$

$$= \frac{\text{ب مج } (س - \bar{س})(ص - \bar{ص})}{\text{مج } (ص - \bar{ص})^2} \quad \dots (٤٨)$$

$$= \frac{\text{ب (مج س ص - مج س } \times \text{ مج ص / ن)}}{\text{مج ص}^2 - (\text{مج ص})^2 / \text{ن}} \quad \dots (٤٩)$$

من ناحية أخرى ، فإن معامل التحديد يبين نسبة تأثير المتغير المستقل (س) على المتغير التابع (ص) أما الجذر التربيعي لمعامل التحديد فيساوي معامل الارتباط المعروف (ر).

وإذا كان المقدار (r^2) يسمى بمعامل التحديد فإن المقدار $(1 - r^2)$ يسمى بمعامل عدم التحديد ، وبديهي فإنه يعبر عن نسبة تأثير العوامل العشوائية على المتغير التابع . وقيمة معامل التحديد (r^2) تتراوح بين الصفر والواحد الصحيح حيث نجد أن :

- ١- إذا كانت $r^2 = \text{صفر}$ ، فهذا يعني أن بسط معامل التحديد يساوى الصفر أى أن التغير المفسر = صفر أى أن التغيرات الكلية فى التابع (ص) ترجع كلها إلى العوامل العشوائية أى أن نسبة تأثير المتغير المستقل س = صفر .
- ٢- إذا كانت $r^2 = ١$ ، فهذا يعني أن التغير الكلى فى التابع (ص) ترجع كلها إلى وجود المتغير المستقل (س) وأن التغيرات العشوائية ذات تأثير منعدم .
- ٣- أما إذا كانت r^2 تقع بين الصفر والواحد الصحيح ، فهذا يعني أن جزءاً من التغير الكلى فى التابع (ص) يرجع إلى وجود المتغير المستقل (س) وجزءاً

آخر يرجع إلى العوامل العشوائية ، وبالطبع كلما زادت قيمة r^2 واقتربت من الواحد الصحيح كان هذا دليلاً على تعاضد تأثير (س) على المتغير التابع والعكس صحيح .

مثال (١١) :

استخدم بيانات مثال رقم (٣) في إيجاد قيمة معامل التحديد :

العل :

$$r^2 = \frac{\text{ب (مج س ص - مج س } \times \text{ مج ص / ن) (الطريقة المباشرة)}}{\text{مج ص}^2 - (\text{مج ص})^2 / \text{ن}}$$

$$\text{أو} \quad r^2 = \frac{\text{ب (مج ح }^2 \text{ - مج ح } \times \text{ مج ح / ن) (الطريقة المختصرة)}}{\text{مج ح}^2 - (\text{مج ح})^2 / \text{ن}}$$

بما أن معادلة خط الانحدار في مثال (٣) كانت على الصورة : ص = ٢٧,١٢ + ١,٦٦ س أى أن ميل خط الانحدار ب = ١,٦٦ . من جدول (٥) نجد أن :

$$r^2 = \frac{1,66 (11216 - 180 \times 10570) - 10570^2 / 10}{10570^2 - 34124} = 0,9712$$

$$0,9712 = \frac{10586,96}{1634} = \frac{906 \times 1,66}{1634} =$$

ومن جدول (٢) نجد أن :

$$r^2 = \frac{906 \times 1,66}{1634} = \frac{[10586,96 - 816] \times 1,66}{10570^2 - 2124} = 0,9712$$

أى أن هناك تطابق تام بين نتائج الطريقتين المباشرة والمختصرة .

نسبة تأثير المتغير المستقل (السماد في هذا المثال) على كمية الإنتاج تبلغ ٩٧,١٢% ، والباقي ٢,٨٨% ترجع إلى مجموعة عوامل عشوائية . أما الجذر التربيعي لمعامل التحديد $= 0.97127 = 0.985$ فهو يمثل قيمة معامل الارتباط بين الظاهرتين (س،ص) وهى نفس النتيجة التى توصلنا إليها فى مثال (٩) .

ملحوظة :

يمكن استخدام الصيغة رقم (٤٧) لحساب معامل التحديد لكنها تمثل مجهود حسابي ضخم ومرهق .

(٨) معادلة خط الانحدار من بيانات مبوبة :

لاحظنا فى الأمثلة التوضيحية السابقة أننا نتعامل مع عينات ذات أحجام صغيرة (ن أقل من ٣٠) ، وذلك بغرض توضيح كيفية استخدام المقاييس التى قدمناها ، لكن صغر حجم العينة يجعلنا تحت رحمة أى خطأ يحدث بالمصادفة عند رصد قيمة أحد المتغيرات ، ناهيك عن أن أى خطأ يحدث يترتب عليه خطأ كبير فى النتائج التى نتوصل إليها لذلك - وكما أكدنا من قبل - يجب ألا يقل عدد الحالات التى نتعامل معها عن ٣٠ حالة (أى أن ن أكبر من ٣٠ وتسمى عينات كبيرة) حتى تكون هناك فرصة لتعادل أثر الحالات الشاذة مع بعضها . ولكن إذا زاد عدد الحالات كثيراً ، فلاشك أن العمل الحسابي سيكون مضاعفاً ومرهقاً إذا ما استخدمنا أسلوب العمل السابق فى حالة البيانات المفردة ، لذا علينا بتحويل هذه البيانات الكبيرة الحجم من حالة بيانات مفردة إلى حالة بيانات مبوبة ، كما وضعنا ذلك - فى مرحلة سابقة - عند الحديث عن معامل الارتباط فى حالة البيانات المبوبة .

والآن نتناول كيفية تقدير معادلتى خط الانحدار ص/س ، س/ص من جدول تكرارى مزدوج وخير وسيلة للتوضيح هى مثال عملي .

مثال (١٣):

- الجدول التالي يوضح توزيع عينة من الأزواج (س) والزوجات (ص) عند بداية سن الزواج لكل منهم ، والمطلوب :
- ١- تقدير عمر الزوج عندما يكون عمر الزوجة ٣٠ سنة .
 - ٢- تقدير عمر الزوجة عندما يكون عمر الزوج ٤٥ سنة .
 - ٣- قياس معامل الارتباط بين عمر الزوج وعمر الزوجة .

جدول (٦)

مراكز فئات س	مراكز فئات ص		١٧	١٩	٢١	٢٣	٢٥	المجموع
	من	من						
٢٠	-١٨	٢	١٦	-١٨	-٢٠	-٢٢	-٢٤	٥
٢٤	-٢٢	٣	٥	٥			٢	١٠
٢٨	-٢٦	١	٤	١٢	٥			٢٢
٣٢	-٣٠		٦	٨		٣		١٧
٣٦	-٣٤		٣	٥	٤			١٢
٤٠	-٣٨						٤	٤
	المجموع	٦	١٨	٢٥	١٢	٩	٧٠	

الحل :

- ١- لتقدير عمر الزوج (س) عندما يكون عمر الزوجة ص = ٣٠ سنة ، يقتضي إيجاد معادلة خط انحدار س على ص ، حيث $S = ج + د ص$. نذكر انه في حالة البيانات المفردة، كنا نجد أن كل قيمة من (س) تتناظرها قيمة من (ص) ، غير أن الأمر يختلف في حالة البيانات المبوبة ، حيث نجد أن كل فئة من (س) يتناظرها العديد من التكرارات الواقعة أمام فئات مختلفة

من (ص) والعكس ، كل فئة من (ص) يناظرها عدد من التكرارات تقع أمام فئات مختلفة من (س) .

ولإيجاد معادلة خط انحدار س على ص من جدول مزدوج ، نقوم بحساب متوسطات قيم (س) المناظرة لكل فئة من فئات (ص) المختلفة ، أى أنه فى معادلة خط انحدار س / ص :

س = ج + د ص ، نعتبر المتغير المستقل (ص) وكأنه مراكز فئات (ص) أما المتغير التابع (س) يؤخذ على أنه متوسط قيم س المناظرة لمراكز فئات (ص) السابقة .

مركز الفئة الأولى للمتغير ص = ١٧

∴ متوسط قيم (س) المقابلة لهذا المركز عبارة عن متوسط حاصل ضرب التكرار المشترك ك_{١١} (والتي تقع أسفل العمود الأول) فى مراكز فئات (س) .

$$\therefore \text{متوسط قيم س} = \frac{١٤٠}{٦} = \frac{٢٨ \times ١ + ٢٤ \times ٣ + ٢٠ \times ٢}{١ + ٣ + ٢} = ٢٣,٣٣$$

مركز الفئة الثانية للمتغير ص = ١٩

∴ متوسط قيم س للمقابلة لهذا المركز =

$$٢٩,٥٥ = \frac{٥٣٢}{١٨} = \frac{٣٦ \times ٣ + ٣٢ \times ٦ + ٢٨ \times ٤ + ٢٤ \times ٥}{٣ + ٦ + ٤ + ٥}$$

مركز الفئة الثالثة للمتغير (ص) = ٢١

$$٣٠,٨٨ = \frac{٧٧٢}{٢٥} = \frac{٣٦ \times ٥ + ٣٢ \times ٨ + ٢٨ \times ١٢}{٥ + ٨ + ١٢}$$

مركز الفئة الرابعة للمتغير (ص) = ٢٣

$$٢٨,٦٧ = \frac{٣٤٤}{١٢} = \frac{٣٦ \times ٤ + ٢٨ \times ٥ + ٢٠ \times ٣}{٤ + ٥ + ٣}$$

مركز الفئة الأخيرة للمتغير (ص) = ٢٥

$$\text{متوسط قيم س المناظرة} = \frac{٤٠ \times ٤ + ٣٢ \times ٣ + ٢٤ \times ٢}{٤ + ٣ + ٢} = \frac{٣٠٤}{٩} = ٣٣,٧٨$$

وبهذه الطريقة نكون قد كونا عمودين : الأول عبارة عن مراكز فئات

ص (وهو المتغير المستقل في معادلة خط انحدار س/ص) والعمود الثاني عبارة عن متوسطات قيم س (وهو المتغير التابع هنا) ومن ثم تسهل عملية إيجاد

معادلة خط الانحدار س = جـ + د ص .

جـ (٧)

ص	س	ص	س
٢٨٩	٣٩٦,٦١	١٧	٢٣,٣٣
٣٦١	٥٦١,٤٥	١٩	٢٩,٥٥
٤٤١	٦٤٨,٤٨	٢١	٣٠,٨٨
٥٢٩	٦٥٩,٤١	٢٣	٢٨,٦٧
٦٢٥	٨٤٤,٥٠	٢٥	٣٣,٧٨
٢٢٤٥	٣١١٠,٤٥	١٠٥	١٤٦,٢١

$$\text{س} = \text{جـ} + \text{د ص} \quad ١٠٥ \times ١٤٦,٢١$$

$$\text{د} = \frac{\text{مجم س ص} - \text{مجم س} \times \text{مجم ص} / \text{ن}}{\text{مجم ص}^2 - (\text{مجم ص})^2 / \text{ن}} = \frac{٣١١٠,٤٥ - \frac{١٠٥ \times ١٤٦,٢١}{٥}}{\frac{١٠٥^2}{٥} - \frac{٢٢٤٥^2}{٥}}$$

$$= \frac{٣٠٧٠,٤١ - ٣١١٠,٤٥}{٢٢٠٥ - ٢٢٤٥} = \frac{٤٠,٠٤٥}{٤٠} = ١,٠٠١$$

$$\text{جـ} = \text{س} - \text{د ص} = \frac{١٤٦,٢١}{٥} - ١,٠٠١ \times \frac{١٠٥}{٥}$$

$$= ٢٩,٢٤ - ٢١,٠٢ = ٨,٢٢$$

$$\therefore \text{س} = ٨,٢٢ + ١,٠٠١ \text{ ص}$$

وعلى ذلك يمكن تقدير عمر الزوج (س) عندما يكون عمر الزوجة ص = ٣٠ سنة على النحو التالي : س = ٨,٢٢ + ١,٠٠١ × ٣٠ = ٣٠,٠٣ + ٨,٢٢ = ٣٨,٢٥ سنة .

٢- لتقدير عمر الزوجة (ص) عندما يكون عمر الزوج س = ٤٥ سنة ، يقتضي إيجاد معادلة خط انحدار ص على س ، (ص = أ + ب س) وبنفس الطريقة السابقة نعتبر (س) هنا وكأنها مراكز فئات المتغير المستقل (س) أما (ص) وهي المتغير التابع فتؤخذ على أنها متوسطات قيم (ص) المناظرة لمراكز فئات س السابقة .

مركز الفئة الأولى للمتغير س = ٢٠

$$\text{متوسط قيم ص المناظرة له} = \frac{٢٣ \times ٣ + ١٧ \times ٢}{٣ + ٢} = \frac{١٠٣}{٥} = ٢٠,٦$$

مركز الفئة الثانية للمتغير س = ٢٤

$$\text{متوسط قيم ص المناظرة} = \frac{٢٥ \times ٢ + ١٩ \times ٥ + ١٧ \times ٣}{٢ + ٥ + ٣} = \frac{١٩٦}{١٠} = ١٩,٦$$

مركز الفئة الثالثة للمتغير س = ٢٨

$$\text{متوسط قيم ص المناظرة} = \frac{٢٣ \times ٥ + ٢١ \times ١٢ + ١٩ \times ٤ + ١٧ \times ١}{٥ + ١٢ + ٤ + ١} = \frac{٤٦٠}{٢٢} = ٢٠,٩$$

مركز الفئة الرابعة للمتغير س = ٣٢

$$\text{متوسط قيم ص المناظرة} = \frac{٢٥ \times ٣ + ٢١ \times ٨ + ١٩ \times ٦}{٣ + ٨ + ٦} = \frac{٣٥٧}{١٧} = ٢١$$

مركز الفئة الخامسة للمتغير س = ٣٦

$$\text{متوسط قيم ص المناظرة} = \frac{٢٣ \times ٤ + ٢١ \times ٥ + ١٩ \times ٣}{٤ + ٥ + ٣} = \frac{٢٥٤}{١٢} = ٢١,١٧$$

مركز الفئة الأخيرة = ٤٠

$$\text{متوسط قيم ص المناظرة} = \frac{٢٥ \times ٤}{٤} = ٢٥$$

ومن الجدول التالي يمكن تقدير معادلة خط انحدار ص على س :

جدول (٨)

س	ص	س ص	س ^٢
٢٠	٢٠,٦	٤١٢	٤٠٠
٢٤	١٩,٦	٤٧٠,٤	٥٧٦
٢٨	٢٠,٩	٥٨٥,٢	٧٨٤
٣٢	٢١	٦٧٢	١٠٢٤
٣٦	٢١,١٧	٧٦٢,١٢	١٢٩٦
٤٠	٢٥	١٠٠٠	١٦٠٠
١٨٠	١٢٨,٢٧	٣٩٠١,٧٢	٥٦٨٠

ص = أ + ب س

$$\therefore \text{ب} = \frac{\text{مجم ص} - \text{مجم س} \times \text{مجم ص} / \text{ن}}{\text{مجم س}^2 - (\text{مجم س})^2 / \text{ن}}$$

$$= \frac{٦/١٢٨,٢٧ \times ١٨٠ - ٣٩٠١,٧٢}{٦/(١٨٠) - ٥٦٨٠} = ٠,١٩١٥$$

$$\text{أ} = \text{ص} - \text{ب س} = \frac{١٢٨,٢٧}{٦} - ٠,١٩١٥ \times \frac{١٨٠}{٦}$$

$$= ٢١,٣٨ - ٥,٧٥ = ١٥,٦٣$$

$$\therefore \text{ص} = ١٥,٦٣ + ٠,١٩١٥ \text{ س}$$

وعلى ذلك فعمر الزوجة (ص) عندما يكون عمر الزوج (س) ٤٥ سنة
من هذه البيانات هو : ص = ١٥,٦٣ + ٠,١٩١٥ × ٤٥ = ١٥,٦٣ + ٨,٦٢ = ٢٤,٢٥ سنة

٣- قياس معامل الارتباط بين عمر الزوج وعمر الزوجة :
من الواضح أنه من الأسهل استخدام العلاقة بين ميل خطي الانحدار
في استنتاج قيمة معامل الارتباط ، بدلاً من إيجاد معامل الارتباط من جدول
مزوج بالطريقة التقليدية .

$$\begin{aligned} \text{.. ص} &= ١٥,٦٣ + ٠,١٩١٥ \text{ س} & \text{.. (انحدار ص/س)} \\ \text{س} &= ١,٠٠١ + ٨,٢٢ \text{ ص} & \text{.. (انحدار س/ص)} \\ \text{.. ر} &= \frac{\text{ب} \times \text{د}}{\text{ج} \times \text{هـ}} & \\ \text{..} &= \frac{٠,١٩١٥ \times ٨,٢٢}{١,٠٠١ \times ٠,٤٤} & \end{aligned}$$

(٩) الاستدلال الإحصائي عن معالم خط الانحدار

بيننا في مرحلة سابقة أن التقديرات واختبارات الفروض الإحصائية ،
يشكلها مع الأدوات الأساسية للاستدلال أو الاستنتاج الإحصائي . نظرية التقدير
تتكون من جزئين : التقدير بنقطة والتقدير بفترة ثقة . اختبارات الفروض
الإحصائية تتكون من : اختبارات معلمية واختبارات لا معلمية . وسوف نركز
في هذا الفصل على التقدير بفترة ثقة وعلى الاختبارات المعلمية لما لهما من
مجال عملي واسع في تحليل الانحدار .

أولاً : التقدير بفترة ثقة :

بفرض أنه من عينة عشوائية توصلنا إلى معادلة خط الانحدار التالية :
ص = ٣ + ٥ س ، أي أن : أ = ٥ ، ب = ٣ وتسمى تلك التقديرات بالتقدير
بنقطة Point Estimate للقيم الحقيقية في المجتمع γ ، β والسؤال هنا : إلى
أي مدى يمكن الاعتماد على تلك التقديرات أ=٥ ، ب=٣ في عملية التنبؤ

بالمتغير التابع ص؟ من المعلوم أنه وبسبب استخدام أسلوب العينة ، ستختلف التقديرات أ ، ب من عينة إلى أخرى ، كما أنها ستختلف عن القيم الحقيقية γ ، β . لكن بتعدد العينات التي يمكن سحبها من المجتمع ، سنجد أن متوسط تلك التقديرات (أي القيمة المتوقعة لها) ، لابد وأن تساوي القيمة الحقيقية في المجتمع ، أي توقع (أ) $\gamma =$ ، توقع (ب) $\beta =$. دعنا نعيد صياغة السؤال السابق بطريقة أخرى : إلى أي مدى تقترب التقديرات أ ، ب من القيم الحقيقية المجهولة γ ، β ؟ للإجابة على ذلك ، فإننا نحتاج إلى قياس الخطأ المعياري لتلك التقديرات حتى يمكننا بناء فترة ثقة للمعالم الحقيقية γ ، β . لكن كيف تنشأ فترة الثقة؟ هذا يتوقف على معرفة التوزيع الاحتمالي للتقديرات أ ، ب . تحت شرط خضوع الحد العشوائي ψ للتوزيع الطبيعي بمتوسط وبتباين على الصورة :

$$(1) \text{ توقع (أ) } \gamma = \text{تباين (أ) } \sigma^2_{\gamma} = \frac{\text{مج-س}^2}{n} \times \frac{1}{(س-س)^2}$$

أي أن : $\gamma \sim A[\gamma, \sigma^2_{\gamma}]$ ، م : تشير إلى التوزيع الطبيعي ، σ^2_{γ} هو تباين الحد العشوائي ψ وهو غالباً مجهول القيمة .

$$(2) \text{ توقع (ب) } \beta = \text{تباين (ب) } \sigma^2_{\beta} = \frac{1}{\text{مج-س}^2} \times \frac{1}{(س-س)^2}$$

أي أن : $\beta \sim B[\beta, \sigma^2_{\beta}]$

وسوف نقتصر على الاستدلال الإحصائي الخاص بالمعلمة (β) فقط لما لها من أهمية خاصة في تحليل الانحدار أما الاستدلال حول (γ) فنأدرأ ما يعول عليه .

بتحويل ب إلى قيمة معيارية ي ، حيث : $y = \frac{\beta - \beta}{\sigma_{\beta}}$ نجد أن ي تتبع توزيع طبيعي ومن ثم يمكن إجراء اختبارات الفروض الإحصائية وإنشاء فترات ثقة للمعلمة β . لكن هذا يتوقف

على معلومية $\hat{\beta}_6$ (تباين الحد العشوائي ψ) وهو قيمة مجهولة ، لذا فإننا نستبدلها بتقدير غير متحيز من عينة عشوائية هو $\hat{\sigma}^2$ وفي ظل استخدام عينة صغيرة الحجم ، يتحول المتغير $\hat{\sigma}^2$ إلى المتغير t ، حيث :

$$t = \frac{\hat{\beta} - \beta}{\hat{\sigma}(\hat{\beta})} \text{ حيث } \hat{\sigma}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\text{مجم (س - س)}} \times \frac{\chi^2_{\alpha/2, n-2}}{n-2} = \frac{\text{مجم (س - س)}}{n-2} \times \frac{\chi^2_{\alpha/2, n-2}}{\text{مجم (س - س)}} \times \frac{\hat{\sigma}^2}{n-2}$$

ع (ب) : الخطأ المعياري للتقدير (ب)

غير : الخطأ المعياري لتقديرات معادلة خط الانحدار .

والمتغير t هو متغير يتبع توزيع t بدرجات حرية $(n-2)$ ، وعلى ذلك يمكن إنشاء فترة ثقة للمعلمة β عند درجة ثقة $(1-\alpha)\%$ على النحو التالي :

(٥٠)...

$$\hat{\beta} \pm t_{(\alpha/2, n-2)} \times \hat{\sigma}(\hat{\beta})$$

مثال (١٣) :

مستخدماً بيانات الأمثلة (٣ ، ٨) قدر β بفترة ثقة ٩٥%

الحل :

من مثال (٣) توصلنا إلى : $\hat{\beta} = 27.12 + 1.66 \text{ س}$ ، $n = 10$

ومن مثال (٨) توصلنا إلى : $\hat{\sigma} = 2.43$ ، وأما الكمية :

$$\hat{\sigma}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\text{مجم (س - س)}} \times \frac{\chi^2_{\alpha/2, n-2}}{n-2} = \frac{2.43^2}{10-2} \times \frac{20.09}{8} = 7.43$$

$$\hat{\beta} \pm t_{(\alpha/2, n-2)} \times \hat{\sigma}(\hat{\beta}) = 1.66 \pm 2.306 \times 2.43 = 1.66 \pm 5.61 = 7.27 \text{ إلى } 7.27$$

$$1.66 \pm 5.61 = 7.27 \text{ إلى } 7.27$$

.. عند درجة ثقة ٩٥% ، نجد أنه في المدى الطويل ، أنه في كل ١٠٠ فترة ثقة للمعلمة β ، نجد أن هناك ٩٥ فترة تحتوي كل منها على المعلمة الحقيقية β ، ومن الخطأ القول بأن هناك احتمال قدرة ٠,٩٥ ، بأن فترة الثقة (١,٤٣) إلى (١,٨٩) تحتوي على المعلمة الحقيقية β .

ثانياً : اختبارات الفروض الإحصائية :

تهتم اختبارات الفروض الإحصائية ببناء قاعدة أو إجراءات ، لاتخاذ قرار بشأن بقبول أو رفض الفرض العدمي ، وهناك أسلوبين لأداء تلك الاختبارات :

أ - فترات الثقة ب- اختبارات المعنوية

أ - اختبارات الفروض : أسلوب فترات الثقة :

إذا وقعت قيمة المعلمة β المراد اختبارها في ظل الفرض العدمي داخل فترة الثقة $(1-\alpha)\%$ فإننا نقبل الفرض العدمي ، وإذا وقعت تلك القيمة خارج حدي الثقة ، فإننا نرفض الفرض العدمي ، فمثلاً : إذا كانت قيمة β المراد اختبارها هي $\beta = ٠,٣$ (أي أن الفرض العدمي هو $\beta = ٠,٣$) وكانت فترة الثقة للمعلمة β على الصورة $\beta = ٠,٤٣$ ، $٠,٥٩$ ، فمن الملاحظ وقوع قيمة β المراد اختبارها وهي $٠,٣$ خارج حدي الثقة ، مما يعني رفض الفرض العدمي ، لأن احتمال مشاهدة قيمة للمعلمة $\beta = ٠,٣$ هو احتمال ضئيل جداً قدره $(\alpha/2)\%$ أما إذا كانت القيمة المراد اختبارها للمعلمة β هي الصفر (أي أن الفرض العدمي هو $\beta = ٠$ صفر) فإننا نبحث عن وجود أو عدم وجود الصفر داخل حدي الثقة ، فإن وقع الصفر داخل حدي الثقة لكان ذلك دليلاً على قبول الفرض العدمي .

ب- اختبارات الفروض : أسلوب اختبار المعنوية :

الأسلوب البديل لفترة الثقة عند اختبار قيمة المعلمة β ، هو أسلوب اختبار المعنوية للمعلمة β . الفكرة الأساسية وراء اختبار المعنوية ، تعتمد على وسيلة الاختبار الإحصائي والتوزيع الاحتمالي لهذه الوسيلة ، ويتم الوصول إلى قرار بقبول أو رفض الفرض العدمي ، بعد مقارنة القيمة العددية لوسيلة الاختبار مع قيمة جدولية تتبع نفس التوزيع الاحتمالي لوسيلة الاختبار .
 نبينا من قبل أن القيمة المعيارية للإحصاء (ب) كانت على الصورة :

$$ت = \frac{(\beta - \beta_0) \sqrt{مجم (س - س_0)}}{ع (ب)} \quad \dots (٥١)$$

ت هنا تعتبر وسيلة الاختبار الإحصائي ، وهي متغير عشوائي يتبع توزيع ت بدرجات حرية (ن-٢) . عند مقارنة القيمة العددية لوسيلة الاختبار (ت) مع قيمة جدولية مستخرجة من توزيع ت عند درجات حرية (ن-٢) ومستوى معنوية $(\alpha / ٢)$ نصل إلى قرار بقبول أو رفض الفرض العدمي .

مثال (١٤) :

مستخدماً بيانات مثال (٣) ، المطلوب اختبار صلاحية معادلة خط الانحدار للاستخدام في عملية التنبؤ عند مستوى معنوية ٥% ، مستخدماً في ذلك أسلوب فترة الثقة واختبار المعنوية.

الحل :

اختبار صلاحية المعادلة للاستخدام تعني إجراء اختبار $\beta = \text{صفر}$.

أ - اختبار $\beta = \text{صفر}$ باستخدام فترة الثقة .

خطوات الاختبار :

١- الفرض العدمي : $\beta = \text{صفر}$. (لا توجد علاقة حقيقية بين س ، ص)

- ٢- الفرض البديل : $\beta \neq \text{صفر}$ (توجد علاقة حقيقية بين س ، ص)
 ٣- فترة الثقة للمعلمة β سبق الحصول عليها في مثال (١٣) وكانت على الصورة : $\beta = 1.43$ إلى 1.89 .
 ٤- وحيث أن قيمة المعلمة المراد اختبارها هي $\beta = \text{صفر}$ ، وحيث أن الصفر لا يقع داخل فترة الثقة فإن القرار هو :
 ٥- القرار : رفض الفرض العدمي وقبول الفرض البديل
 ∴ $\beta \neq \text{صفر}$ ، أي هناك علاقة ارتباط حقيقية بين المتغيرين (س ، ص)

ب- اختبار $\beta = \text{صفر}$ باستخدام اختبار المعنوية (اختبار ت) :

خطوات الاختبار :

- ١- الفرض العدمي : $\beta = \text{صفر}$.
 ٢- الفرض البديل : $\beta \neq \text{صفر}$
 ٣- ت = $\frac{(\text{ب} - \beta)}{\text{ع (ب)}} = \frac{(\text{ب} - \text{صفر})}{\frac{\text{مجم س}^2 - (\text{مجم س})^2 / \text{ن}}{24 \div 2.43}} = \frac{1.66}{16.395}$
 ٤- ت الجدولية : ت (ن - ٢ ، $\alpha / 2$) = ت (٨ ، ٠.٠٢٥) = ٢.٣٠٦
 ٥- المقارنة : ت المحسوبة أكبر من ت الجدولية .
 ٦- القرار : رفض الفرض العدمي ومن ثم قبول الفرض البديل
 ∴ $\beta \neq \text{صفر}$ وهو نفس القرار الذي توصلنا إليه باستخدام أسلوب فترة الثقة .

(١٠) تحليل التباين وتحليل الانحدار :

يمكن استخدام أسلوب تحليل التباين في اختبار مدى وجود علاقة حقيقية بين المتغير التابع (ص) والمتغير المستقل (س) ، أي في اختبار $\beta = \text{صفر}$ ، وذلك بتجزئة الاختلافات الكلية في المتغير التابع (ص) إلى مركبتين : الأولى ترجع إلى اختلافات ناتجة عن العلاقة الانحدارية بين (ص ، س) أي اختلافات تفسر بسبب وجود المتغير المستقل (س) ، والثانية ترجع إلى

اختلافات ناتجة عن مجموعة من المتغيرات العشوائية ، أي اختلافات لا يمكن ارجاعها أو تفسيرها إلى سبب معين ، أي اختلافات غير مفسرة ، وقد بينا عند الحديث عن معامل التحديد ما يلي :

التغير الكلي في المتغير التابع (ص) = مجموع المربعات الكلي (م.م.ك)

$$= \text{م.ج (ص-ص)}^2 = \text{م.ج ص}^2 - (\text{م.ج ص})^2 / \text{ن}$$

التغير المفسر بسبب وجود المتغير المستقل = مجموع المربعات بسبب الانحدار (م.م.ب)

$$= \text{م.ج (ص-ص)}^2$$

$$= \text{ب.م.ج (س-س)}^2 / (\text{ص-ص})$$

$$= \text{ب.م.ج (س-ص)}^2 / (\text{ص-ص})$$

التغيير غير المفسر والذي يرجع إلى عوامل عشوائية = مجموع مربعات

$$\text{البواقي (م.م.د)} = \text{م.ج (ص-ص)}^2$$

$$\text{م.م.ك} = \text{م.م.ب} + \text{م.م.د}$$

وباسترجاع أسلوب تحليل التباين من الباب السابق ، نجد أن كل مجموع مربعات يرتبط دائما بدرجات حرية (د. ح) معينة ، فمثلا مجموع المربعات الكلي (م.م.ك) يرتبط بعدد (ن-1) من درجات الحرية ، (تم خصم درجة حرية واحدة من حجم العينة مقابل تقدير ص) ، بينما مجموع مربعات البواقي (م.م.د) قلته (ن-2) من درجات الحرية (تم خصم درجتين حرية من حجم العينة مقابل تقدير المعلمتين γ ، β) ، وحيث أن :

$$\text{د. ح (م.م.ك)} = \text{د. ح (م.م.ب)} + \text{د. ح (م.م.د)}$$

$$\text{ن-1} = \text{د. ح (م.م.ب)} + (\text{ن-2})$$

∴ درجات الحرية لمجموع المربعات بسبب علاقة الانحدار تساوي واحد (وهي في الواقع تساوي عدد المتغيرات المستقلة في نموذج الانحدار) .
ويترتب مجاميع المربعات السابقة مقترنة بدرجات الحرية في جدول يسمى جدول تحليل التباين Analysis of Variance نصل إلى الصورة التالية .

جدول تحليل التباين

مصدر الاختلاف	مجموع المربعات (م.م)	درجات الحرية	متوسط مجموع المربعات (م.م.م)
رجعا إلى الانحدار (م.م.ب)	مجـ (ص - ص) ^٢ = ب مجـ (س - س) × (ص - ص)	١	مجـ (ص - ص) ^٢ / ١
رجعا إلى البواقي م.م.د	مجـ (ص - ص) ^٢	ن - ٢	مجـ (ص - ص) ^٢ / (ن - ٢) = ص ^٢ خ
م . م . ك	مجـ (ص - ص) ^٢	ن - ١	

ولاختبار الفرض العدمي القائل بأن $\beta = 0$ صفر ، أي لا توجد علاقة خطية بين ص ، س أي لا يوجد تأثير للمتغير المستقل س على المتغير التابع ص ، فإننا نقارن قيمة متوسط مجموع المربعات بسبب الانحدار مع قيمة متوسط مجموع المربعات للبواقي ، حيث يكون خارج القسمة لهما هو متغير عشوائي يتبع توزيع ف :

$$ف = \frac{\text{مجـ (ص - ص)}^2}{\text{مجـ (ص - ص)}^2 / (ن - ٢)} = \frac{\text{ب مجـ (س - س) (ص - ص)}}{\text{ص}^2 \text{ خ}} \dots (٥٢)$$

حيث $\text{ص}^2 \text{ خ}$ هو تباين تقدير معادلة خط الانحدار أو تباين البواقي ، وبمقارنة القيمة العددية للمتغير ف مع قيمة جدولية مستخرجة من جدول توزيع ف عند درجات حرية (١ ، ن - ٢) ومستوى معنوية $(\alpha / ٢)$ ، يمكن اتخاذ قرار بقبول أو رفض الفرض العدمي . بالطبع - وكما هو متوقع - تكون نتيجة القرار هنا هي نفس النتيجة التي نصل لها سواء استخدمنا اختبار معنوية β باستخدام توزيع (ت) أو باستخدام فترة الثقة للمعلمة β .

ملحوظة :

- ١- عند ثبات درجات الحرية ومستوى المعنوية ، فإن مربع قيمة ت الجدولية = قيمة ف الجدولية .
- ٢- من الممكن أن نصل إلى صيغة أخرى لوسيلة الاختبار ف بدلالة معامل التحديد ^٢ على النحو التالي :

$$\frac{r_1(\bar{m} - m)}{r_2(\bar{m} - m)} = \frac{r_1(\bar{m} - m)}{r_2(\bar{m} - m)} \quad \& \quad \frac{r_1(\bar{m} - m)}{r_2(\bar{m} - m)} = \frac{r_1(\bar{m} - m)}{r_2(\bar{m} - m)}$$

ج۔ ف۔ مج۔ (ص^۱ - ص^۲) × (ن - ۲)

مجـ (ص - ص^٨)
بقسمة كل من البسط والمقام على مجـ (ص - ص) نصل إلى :

(०२)...

$$\text{ف} = \frac{r}{r-1} \times (n-2)$$

أي أنه إذا علمت قيمة r^2 ، أمكن الوصول إلى قيمة وسيلة الاختبار (ف) والعكس صحيح .

مثال (۱۵)

بفرض أنك حصلت على البيانات التالية :

٦	٥	٤	٣	٢	١	س
١٠	٩	٦	٤	٤	٢	ص

المطلوب :

- ١- معادلة خط انحدار ص / س
٢- اختبار معنوية معادلة خط الانحدار أي اختبار β - صفر مستخدما في ذلك أسلوب تحليل التباين عند $\alpha = 0\%$ حيث القيمة الجدولية : ف (١ ، ٤ ، ٧٠١) = ٧٠١ %

٣- معامـل التـحـديـد ومعامـل الارتبـاط .

الحل :

س	ص	س	ص	س
۱	۲	۱	۴	۲
۲	۴	۴	۱۶	۹
۳	۴	۹	۱۶	۱۲
۴	۶	۱۶	۳۶	۲۴
۵	۹	۲۵	۸۱	۴۵
۶	۱۰	۳۶	۱۰۰	۶۰
۲۱	۳۵	۹۱	۲۵۳	۱۵۱

١- معادلة خط انحدار ص / س : ص = أ + ب س ، حيث

$$\begin{aligned} \frac{6/30 \times 21 - 101}{6/2(21) - 91} &= \frac{\text{مـجـسـ صـ - مجـسـ} \times \text{مـجـ صـ} / \text{ن}}{\text{مـجـسـ} - 2 - (\text{مـجـسـ}) / 2 / \text{ن}} = \text{ب} \\ 1,628 &= \frac{28,5}{17,5} = \frac{122,5 - 101}{73,5 - 91} \\ \frac{21}{6} \times 1,628 - \frac{35}{6} &= \frac{\text{مـجـ صـ}}{\text{ن}} \times \text{ب} - \frac{\text{مـجـ صـ}}{\text{ن}} = \text{أ} \\ 0,135 &= 5,698 - 5,833 = \\ \therefore \text{ص} &= 1,628 + 0,135 \end{aligned}$$

٢- اختبار β - صفر عن طريق تحليل التباين

يقتضي جدول تحليل التباين حساب الكميات التالية :

م . م . ك = مجـ ص ٢ - (مجـ ص ٢ / ن - ٢٥٣ - ٦/٢(٣٥)

$$48,83 - 204,17 - 203 =$$

$$م. م. بسبب الانحدار = ب \times (م. م. ص - م. م. س \times م. م. ص / ن) \\ 46,398 = 28,5 \times 1,628 - (122,5 - 101) = 46,398$$

جدول تحليل التباين

المصدر	م. م.	د. ح.	م. م. م.	نسبة التباين ف
م. م. بسبب انحدار	46,398	1	46,398	
ص / س				76,3125
م. م. د (البواقي)	2,432	4	0,608	
م. م. ك	48,830	5		

خطوات الاختبار :

١- الفرض العدمي $\beta = 0$ (لا توجد علاقة حقيقية بين س ، ص)

٢- الفرض البديل $\beta \neq 0$ (توجد علاقة حقيقية بين س ، ص)

٣- ف المحسوبة = 76,3125

٤- ف الجدولية : ف (١ ، ٤ ، ٥) = ٧,٧١

٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .

٦- القرار : رفض الفرض العدمي ومن ثم قبول الفرض البديل

∴ $\beta \neq 0$ أي هناك علاقة معنوية بين س ، ص

$$٣- معامل التحديد ر^2 = \frac{م. م. بسبب الانحدار}{م. م. ك} = \frac{46,398}{48,83} = 0,9502$$

أي أن المتغير المستقل يؤثر في المتغير التابع بنسبة ٩٥% تقريبا ، والباقي يرجع لتأثير العوامل العشوائية أو بمعنى آخر ٩٥% من الاختلافات الكلية في التابع ص ، ترجع إلى وجود المتغير المستقل س ، أما معامل الارتباط فهو

الجذر التربيعي لمعامل التحديد وإشارة معامل الارتباط هي نفس إشارة ميل خط الانحدار (ب) .

... معامل الارتباط $= \sqrt{0.9502} = 0.9747$ ، وحيث أن إشارة ميل خط الانحدار (ب) موجبة ، يكون هناك ارتباط قوي وموجب بين (ص ، س) .

(١١) استخدام نموذج الانحدار في عملية التنبؤ بفترة ثقة :

بعد أن توصلنا إلى وصف العلاقة بين المتغير التابع (ص) والمستقل (س) والتي على الصورة : $\hat{V} = a + bS$ ، وبعد أن تأكدنا من صحة العلاقة بينهما عن طريق اختبار $\beta = 0$ صفر ، يتبقى لنا كيفية استخدام تلك العلاقة الانحدارية في التنبؤ بقيمة المتغير التابع (ص) عند قيمة محددة معلومة للمتغير المستقل (س) ، هنا ينشأ نوعين من التنبؤ :

(١) التنبؤ بمتوسط قيم (ص) عند قيمة محددة معلومة للمتغير س .

(٢) التنبؤ بقيمة (ص) عند قيمة محددة معلومة للمتغير س .

للتوضيح ، في مثال (٣) كانت العلاقة بين كمية الانتاج من القمح (ص) وكمية السماد المستخدم (س) على الصورة : $\hat{V} = 27.12 + 1.66S$. هنا قد نرغب في تقدير متوسط كمية الانتاج لعدد من الألفنة عندما نستخدم في كل فدان ١٥ كيلو سماد مثلاً أي $S = 15$ أو قد نرغب في تقدير انتاج فدان واحد فقط عندما يعطي كمية سماد $S = 15$ وحدة .

لكن أي التقديرين : تقدير متوسط قيم \hat{V} لعدة ألفنة أو تقدير قيمة \hat{V} في فدان واحد ، يحقق دقة أكبر ؟ قبل الإجابة على ذلك ، دعنا أولاً ننتبأ بتلك القيم .

(أ) تقدير متوسط الإنتاج (\hat{V}) لعدة فترات عندما تكون س في كل فترة $= 15$ هي :

$$\hat{V} = 27.12 + 1.66 \times 15 = 52.02 \text{ كيلو جرام}$$

(ب) تقدير كمية الانتاج (\hat{V}) عندما تكون $S = 15$ هي

$$\hat{ص} = ٢٧,١٢ + ١,٦٦ \times ١٥ = ٥٢,٠٢ \text{ كيلو جرام}$$

من الواضح أن كلا التقديرين : متوسط ($\hat{ص}$) ، $\hat{ص}$ هما نفس القيمة عند $س = ١٥$ ، كلا التقديرين يعتبران تقديراً بنقطة أو تقدير وحيد القيمة) ، لكن الفرق بينهما يكمن في الدقة النسبية لكل منهما ، وهذه الدقة النسبية تقاس عن طريق خطأ المعاينة أو الخطأ المعياري لكل تقدير ، وهي على الصورة التالية:

أ- تقدير الخطأ المعياري للمقدر $\hat{ص}$ ، حيث $\hat{ص}$ هي متوسط قيم $\hat{ص}$ عند قيمة محددة معلومة للمتغير $س$ ولكن $س$ هو :

$$ع(\hat{ص}) = \sqrt{\frac{١}{ن} + \frac{(\bar{س} - س)^2}{مج(س-س)}} \quad (٥٤) \dots$$

حيث $ع$ هو الخطأ المعياري لمعادلة خط الانحدار = $\frac{مج(ص-ص)}{ن-٢}$

ب- تقدير الخطأ المعياري للمقدر $\hat{ص}$ ، حيث $\hat{ص}$ هي قيمة $ص$ عند قيمة محددة معلومة للمتغير $س$ ولكن $س$ هو :

$$ع(\hat{ص}) = \sqrt{\frac{١}{ن} + \frac{(\bar{س} - س)^2}{مج(س-س)}} + ١ \quad (٥٥) \dots$$

وعن طريق الأخطاء المعيارية يمكن إنشاء فترة ثقة للتقديرات ($\hat{ص}$) سواء أكان التقدير هو متوسط قيم $\hat{ص}$ أو قيمة $\hat{ص}$ المفردة على النحو التالي :

$$\bar{ص} \pm \hat{ص} \times ت(٢/١٠٠ - \alpha) \times ع(\hat{ص}) \quad (٥٦) \dots$$

حيث : $\hat{ص} = أ + ب س$ ، $ع(\hat{ص})$ إما أن تكون المعادلة (٥٤) في حالة تقدير متوسط $\hat{ص}$ أو تكون المعادلة (٥٥) في حالة تقدير قيمة وحيدة لـ $\hat{ص}$.

واستكمالاً للمثال المطروح بين أيدينا نعلم من هذا المثال أن :

$ن = ١٠$ ، $\bar{س} = ١٠ \div ١٨٠ = ١٨$ ، $ع(ص) = ٢,٤٣$ ، $مج(س-س) = ٥٧٦$ والمطلوب هو :

١- تقدير متوسط $\hat{\mu}$ بفترة ثقة ٩٥% عند س. = ١٥ .

٢- تقدير $\hat{\mu}$ بفترة ثقة ٩٥% عند س. = ١٥ .

بالنسبة للمطلوب الأول : تقدير متوسط $\hat{\mu}$ بفترة ثقة ٩٥% :

$$\hat{\mu} = \bar{x} \pm t_{(n-2, \alpha/2)} \times \frac{s}{\sqrt{n}}$$

$$= (\bar{x} \pm t_{(n-2, \alpha/2)} \times \frac{s}{\sqrt{n}}) \pm \frac{1}{n} \sqrt{\frac{s^2}{n-2} + \frac{1}{n}}$$

$$= (15 \pm 2.073 \times \frac{2.43}{\sqrt{18}}) \pm \frac{1}{18} \sqrt{\frac{(18-1) \times 2.43^2}{17} + 1}$$

$$= 15 \pm 0.826 \times 2.43 \pm 0.052 = 14.91 \pm 0.052 = 14.91 \text{ و } 15.052$$

أي أنه عندما يتم استخدام كمية السماد س = ١٥ وحدة في العديد من

الأقنعة ، فإن متوسط إنتاج الفدان المتوقع سينتج بين ١٤.٩١ وحدة ، ١٥.٠٥٢ وحدة ،

وحدة بدرجة ثقة ٩٥% (لاحظ أن مدى الثقة هنا = ١٤.٩١ - ١٥.٠٥٢ = ٠.١٤٢) .

وبالنسبة للمطلوب الثاني : تقدير قيمة $\hat{\mu}$ بفترة ثقة ٩٥% :

$$\hat{\mu} = \bar{x} \pm t_{(n-2, \alpha/2)} \times \frac{s}{\sqrt{n}}$$

$$= (\bar{x} \pm t_{(n-2, \alpha/2)} \times \frac{s}{\sqrt{n}}) \pm \frac{1}{n} \sqrt{\frac{s^2}{n-2} + \frac{1}{n}}$$

$$= (15 \pm 2.073 \times \frac{2.43}{\sqrt{18}}) \pm \frac{1}{18} \sqrt{\frac{(18-1) \times 2.43^2}{17} + 1}$$

$$= 15 \pm 0.826 \times 2.43 \pm 0.052 = 14.91 \pm 0.052 = 14.91 \text{ و } 15.052$$

أي أنه عند استخدام كمية السماد س = ١٥ وحدة في فدان واحد ، فإن

إنتاجه المتوقع سينتج بين ١٤.٩١ ، ١٥.٠٥٢ وحدة بدرجة ثقة ٩٥%

(لاحظ أن مدى الثقة هنا = ١٤.٩١ - ١٥.٠٥٢ = ٠.١٤٢) .

- من مقارنة فترة الثقة لكلا التقديرين ، يمكن أن نتوصل للنتائج التالية :
- ١- بصفة عامة كلما زاد حجم العينة ، كلما قل الفرق بين حدى الثقة ، وكلما اقترب التقدير من القيمة الحقيقية المجهولة وهذا يعنى زيادة الدقة فى التقدير .
 - ٢- اتساع فترة الثقة عند تقدير $\hat{\mu}$ عن فترة الثقة عند تقدير متوسط $\hat{\mu}$ ، وهذا يعنى أن الخطأ المعياري فى تقدير $\hat{\mu}$ أكبر من الخطأ المعياري فى تقدير متوسط $\hat{\mu}$ ، وهذا واضح إذ أن $\hat{\mu}$ فى الحالة الأولى = ٢,٥٦٧ بينما $\hat{\mu}$ فى حالة المتوسط = ٠,٨٢٦ .
 - ٣- الخطأ المعياري فى تقدير متوسط $\hat{\mu}$ أو فى تقدير $\hat{\mu}$ يصل إلى أدنى قيمة له عندما تكون $\bar{s} = \bar{s}$.
 - ٤- كلما ابتعدت \bar{s} عن \bar{s} كلما زاد الخطأ المعياري فى تقدير $\hat{\mu}$ أو متوسط $\hat{\mu}$ والعكس صحيح ، لذا ينصح أن يتم التنبؤ بقيمة $\hat{\mu}$ عند قيمة \bar{s} القريبة بقدر الإمكان من \bar{s} .
 - ٥- إذا ابتعدت \bar{s} عن \bar{s} بدرجة كافية ووقعت خارج مجال قيم \bar{s} بالعينة ، يكون من الخطر عمل أى استنتاجات حول $\hat{\mu}$ أو متوسط $\hat{\mu}$.

مثال (١٦) (مثال شامل) :

البيانات التالية تبين قيمة الإنفاق على الدعاية بالآلاف جنيه (س) وقيمة المبيعات بالآلاف جنيه (ص) لإحدى الشركات خلال خمسة شهور متتالية:

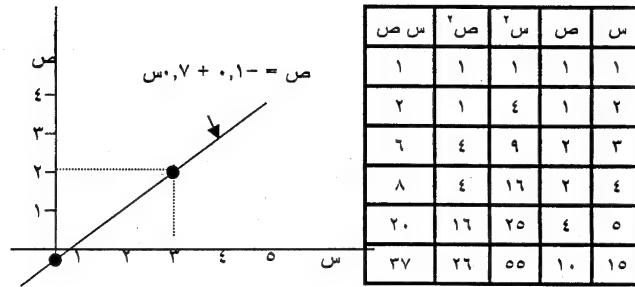
الشهر	يناير	فبراير	مارس	إبريل	مايو
الإنفاق (س)	١	٢	٣	٤	٥
المبيعات (ص)	١	١	٢	٢	٤

المطلوب :

- ١- ارسم الشكل الانتشارى لهذه البيانات ثم وفق لها أفضل خط مستقيم مستخدماً طريقة المربعات الصغرى .

- ٢- أوجد قيمة معامل التحديد ومعامل الارتباط .
- ٣- قدر الخطأ المعياري لنموذج خط الانحدار .
- ٤- قدر بفترة ثقة ٩٥% ميل خط الانحدار β .
- ٥- اختبر صلاحية نموذج خط الانحدار في الحالات التالية :
 - (أ) $\beta = 0$ - صفر
 - (ب) $\beta = 0.3$ ، حيث $\alpha = 0.05$.
- مستخدماً في الحالتين أسلوب فترة الثقة ثم أسلوب اختبار المعنوية .
- ٦- اختبر صلاحية نموذج خط الانحدار ، مستخدماً في ذلك أسلوب تحليل التباين عند مستوى معنوية ٥% .
- ٧- قدر بفترة ثقة ٩٥% ما يلي :
 - أ- متوسط قيمة المبيعات في الشهور القادمة عندما تكون قيمة الدعاية في أى شهر س. = ٤ (ألف جنيه) .
 - ب- قيمة المبيعات في الشهر القادم عندما تكون قيمة الدعاية في ذلك الشهر س. = ٤ (ألف جنيه) .
 - ج- أعد المطلوب (أ) مرة أخرى عند س. = ٣ ثم عند س. = ٧ ثم علق على النتائج.

الحل:



١- الشكل الانتشاري، وأفضل خط مستقيم .

عند رصد البيانات بيانياً نلاحظ أنه يمكن استبدالها تقريباً بخط مستقيم ينتجه صعوداً من أسفل اليسار إلى أعلى اليمين ، ولرسم أفضل خط مستقيم بطريقة المربعات الصغرى نجد أن :

ص^٨ = أ + ب س حيث

$$ب = \frac{\text{مجم س ص} - \text{مجم س} \times \text{مجم ص} / \text{ن}}{\text{مجم س}^2 - (\text{مجم س})^2 / \text{ن}} = \frac{٣٧ - ٥/١٠ \times ١٥}{١٠ - ٥^2 / (١٥)} = ٠,٧$$

(وتسمى ب بمعزل خط الانحدار أو معدل تغير ص مع س عندما تتغير س بوحدة واحدة).

$$أ = \frac{\text{مجم ص}}{\text{ن}} - ب \times \frac{\text{مجم س}}{\text{ن}} = \frac{١٥}{١٠} - ٠,٧ \times \frac{٣٧}{١٠} = ٠,١ - ٢,١ = -٢,٠$$

(وتسمى أ بالجزء المقطوع من المحور الرأسي وهي مقدار ثابت ، أي قيمة ص عندما تكون س=صفر) .

∴ معادلة خط الانحدار تصبح على الصورة : ص^٨ = ٠,١ - ٢,٠ س

ولرسم هذا الخط نكتفي برصد نقطتين بيانياً فمثلاً :

عند س = صفر فإن ص = ٠,١ ∴ أول نقطة (٠,١)

عند س = ٢ فإن ص = ٢ ∴ ثاني نقطة (٢,٢)

∴ عن طريق رصد نقطتين (٠,١) ، (٢,٢) والتوصيل بينهما بخط مستقيم ، نحصل على أفضل خط انحدار يمثل تلك البيانات .

٢- معامل التحديد ومعامل الارتباط :

$$\text{معامل التحديد } R^2 = \frac{\text{التغير المفسر}}{\text{التغير الكلي}} = \frac{\text{مجم (ص}^2 - \text{ص}^{\bar{}}^2)}{\text{مجم (ص}^2 - \text{ص}^{\bar{}}^2)}$$

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{0,7 \times 5 - (0,7 \times 10 - 37)}{5 - 26} = \frac{0,7 \times 5 - 0,7}{-21} = \frac{0,7 \times 4}{-21} = -0,1333$$

أى أن ٨١,٦٦% من التغير الكلى فى المبيعات (ص) ترجع إلى الإنفاق على الدعاية ، وأن الباقي ١٨,٣٤% ترجع إلى عوامل عشوائية ، معامل الارتباط هو الجذر التربيعي لمعامل التحديد أما إشارته (موجب أو سالب) فتحدد وفق إشارة ميل خط الانحدار ب .

$$r = \sqrt{0,81666} = 0,9$$

وحيث أن إشارة (ب) هى موجبة ، تكون إشارة معامل الارتباط هى أيضاً موجبة ، أى أن هناك ارتباط طردى (موجب) قوى بين قيمة الإنفاق على الدعاية وقيمة المبيعات المتحققة .

٣- الخطأ المعياري لنموذج خط الانحدار :

$$s_{\text{خط}} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}} \quad \text{أو} \quad s_{\text{خط}} = \sqrt{\frac{\sum (y_i - \bar{y})^2 - \frac{(\sum (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum (x_i - \bar{x})^2}}{n-2}}$$

ويمكن حساب قيمة خط باى صيغة من تلك الصيغ .

س	ص	ص ^٨ - ص ^٧ + ٠,١ - ص ^٨	ص ^٨ - ص ^٧	(ص ^٨ - ص ^٧) ^٢
١	١	٠,٦	٠,٦ - ١ = -٠,٤	٠,١٦
٢	١	١,٣	١,٣ - ١ = ٠,٣	٠,٠٩
٣	٢	٢	٢ - ٢ = ٠	صفر
٤	٢	٢,٧	٢,٧ - ٢ = ٠,٧	٠,٤٩
٥	٤	٣,٤	٣,٤ - ٤ = -٠,٦	٠,٣٦
			صفر	١,١٠

$$\begin{aligned} \text{غير} &= \frac{\text{مج} - (\text{ص} - \text{ص})}{\text{ن} - 2} = \frac{1,1}{3} \sqrt{\frac{1}{3}} = 0,366 \sqrt{\frac{1}{3}} = 0,61 \\ \text{أو غير} &= \frac{\text{مج} - \frac{(\text{مج} - \text{ص})^2}{\text{ن}}}{\text{ن} - 2} = \frac{(37) \cdot 0,7 - 10 \times (0,1) - 26}{3} \sqrt{\frac{1}{3}} = \frac{25,9 - 1 + 26}{3} \sqrt{\frac{1}{3}} = 0,61 \end{aligned}$$

٤- تقدير ميل خط الانحدار β بفترة ثقة ٩٥% :

$\beta = \pm t_{(n-2, \alpha/2)} \times \text{ع(ب)}$ ، (ع(ب) : الخطأ المعياري للتقدير ب)

$$\beta = \pm t_{(n-2, \alpha/2)} \times \frac{\text{غير}}{\sqrt{\frac{\text{مج} - \frac{(\text{مج} - \text{ص})^2}{\text{ن}}}{\text{ن} - 2}}}$$

$$= \pm 0,7 \times \frac{0,61}{\sqrt{\frac{0,61^2 - 10 \times (0,1)^2 - 26}{3}}}$$

$$= \pm 0,7 \times 3,182 = \pm 2,2274$$

∴ معامل خط الانحدار β تتراوح قيمته بين ٠,٠٩ ، ١,٣١ بدرجة ثقة ٩٥% .

٥- اختبار صلاحية نموذج خط الانحدار :

أولاً : باستخدام أسلوب فترة الثقة :

$$\beta = \pm t_{(n-2, \alpha/2)} \times \text{ع(ب)} = \pm 0,7 \times 3,182 = \pm 2,2274$$

أ- عندما يكون الفرض العدمي هو $\beta = 0$

حيث أن الصفر يقع خارج حدى الثقة ، فهذا يعني رفض الفرض العدمي بأن $\beta = 0$ ومن ثم قبول الفرض البديل أي $\beta \neq 0$ وهذا يعني وجود تأثير معنوي للمتغير المستقل على المتغير التابع .

ب- عندما يكون الفرض العدمي هو $\beta = 0.3$ ، حيث أن القيمة المراد اختبارها للفرض العدمي هو $\beta = 0.3$ تقع داخل حدى الثقة ، فهذا يعني قبول نص الفرض العدمي أى أن $\beta = 0.3$.

ثانياً : باستخدام أسلوب اختبار المعنوية :

أ- عندما يكون الفرض العدمي هو $\beta = 0$ - صفر

خطوات الاختبار :

١- الفرض العدمي : $\beta = 0$ - صفر

٢- الفرض البديل : $\beta \neq 0$ - صفر

$$٣- ت = \frac{\beta - ٠}{\text{ع.ب.}} = \frac{٠.٧ - \text{صفر}}{٠.١٩} = ٣.٦٨٠$$

$$٤- ت الجدولية : ت (٧/٢٠٠) = ت (٧/٢٠٠) = ٣.١٨٢$$

٥- المقارنة: ت المحسوبة أكبر من ت الجدولية .

٦- القرار: رفض الفرض العدمي وقبول الفرض البديل.

∴ $\beta \neq 0$ - صفر وهى نفس النتيجة التى توصلنا إليها عن طريق فترة الثقة.

ب- عندما يكون الفرض العدمي هو $\beta = 0.3$ -

خطوات الاختبار :

١- الفرض العدمي : $\beta = 0.3$ -

٢- الفرض البديل : $\beta \neq 0.3$ -

$$٣- ت = \frac{\beta - ٠.٣}{\text{ع.ب.}} = \frac{٠.٧ - ٠.٣}{٠.١٩} = ٢.١١$$

$$٤- ت الجدولية : ت (٧/٢٠٠) = ت (٧/٢٠٠) = ٣.١٨٢$$

٥- المقارنة : ت المحسوبة أقل من ت الجدولية .

٦- القرار : قبول الفرض العدمى ، وهى نفس النتيجة التى توصلنا إليها عن طريق فترات الثقة.

٦- استخدام تحليل التباين فى تحليل الانحدار :

لتكوين جدول تحليل التباين ، يلزم حساب المجاميع التالية :

$$\begin{aligned} \text{م.م.ك} &= \text{م.م.ص}^2 - (\text{م.م.ص})^2 / \text{ن} = 26 - 5^2 / (10) = 6 \\ \text{م.م.} &\text{بسبب الانحدار} = \text{ب} \times (\text{م.م.ص} - \text{م.م.ص} \times \text{م.م.ص} / \text{ن}) \\ &= 0.7 \times (5 - 5 \times 10 / 10) = 0.7 \times 0.7 = 0.49 \end{aligned}$$

جدول تحليل التباين

المصدر	م.م.	د. ح	م.م.م	ف
م.م. بسبب انحدار ص/س	٠,٤٩	١	٤,٩	
م.م. د. البواقي	١,١	٣	٠,٣٦٦	١٣,٣٨٨
م.م.ك	٦	٤		

خطوات الاختبار :

- ١- الفرض العدمى $\beta = \text{صفر}$
 - ٢- الفرض البديل $\beta \neq \text{صفر}$
 - ٣- $\text{ف} = 13,388$
 - ٤- $\text{ف الجدولية} = \text{ف} (1, 3, 5\%) = 10,13$
 - ٥- المقارنة : $\text{ف المحسوبة أكبر من ف الجدولية}$.
 - ٦- القرار : رفض الفرض العدمى وقبول الفرض البديل .
- ∴ $\beta \neq \text{صفر}$ أى أن هناك تأثير للمتغير المستقل على المتغير التابع .

ملاحظات :

١- القرار الذى نتوصل إليه باستخدام تحليل التباين ، هو نفس القرار الذى نتوصل إليه سواء باستخدام فترة الثقة أو باستخدام اختبار المعنوية للمعامل β عن طريق توزيع ت .

٢- بمعلومية قيمة ت الجدولية نصل إلى قيمة ف الجدولية أو العكس ، حيث مربع قيمة ت الجدولية = قيمة ف الجدولية ، فمثلاً إذا كانت قيمة ف الجدولية = ١٠,١٣ فإن قيمة ت الجدولية $\sqrt{10.13} = 3.182$ وهى نفس القيمة التى استخدمناها من قبل ، هذا على فرض ثبات حجم العينة ومستوى المعنوية .

٣- من جدول تحليل التباين يمكن استنتاج معامل التحديد :

$$r^2 = \frac{\text{م.م بسبب الانحدار}}{\text{م.م ك}} = \frac{4.9}{6} = 0.81666$$

$$\text{ومعامل الارتباط } r = 0.81666$$

٤- من الممكن حساب قيمة ف بمعلومية معامل التحديد وفق العلاقة التالية :

$$F = \frac{\frac{r^2}{1-r^2} \times \frac{n-k}{k}}{\text{حيث ك : عدد المعالم المقدرة وهى هنا :}}$$

$$F = 2.07 \text{ أى أن ك } = 2$$

$$F = \frac{2-0}{1-2} \times \frac{0.8166}{0.1834} = 13.358$$

وهى نفس القيمة التى ظهرت فى جدول تحليل التباين (هناك اختلاف

ضئيل بسبب تقريب العمليات الحسابية) .

٧- التنبؤ بفترة ثقة ٩٥% :

أ- تقدير متوسط المبيعات عند س = ٤ بفترة ثقة ٩٥% .

$$\text{لمر } \frac{1}{\text{س}} = \text{ص}^{\wedge} + \text{ت} (ن-2) \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$= (1 + \text{ب من}) + \text{ت} (0.25, 3) \times \sqrt{\frac{1}{5} + \frac{(\text{س} - \text{س}^{\wedge})^2}{\text{ن} - 2}} \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

من هذا المثال توصلنا إلى :

$$1 - 0.1 = \text{ب} = 0.9, \text{ع} = 0.61, \text{س} = 3$$

$$\text{مجم س}^{\wedge} - \text{مجم س} (\text{مجم س}) / \text{ن} = 55 - 5 / 15 = 10$$

$$\text{لمر } \frac{1}{\text{س}} = \text{ص}^{\wedge} + \text{ت} (ن-2) \times \frac{1}{\text{ع} (\text{ص}^{\wedge})} \times 0.61 \times 3,182 + (4 \times 0.7 + 0.1 - 3) = 2.70$$

$$= 2.70 + 3,182 \times 0.33 = 2.7 + 1.06 = 3.76, 1.64$$

أى أنه عندما تقوم الشركة بإنتاج 4 آلاف جنيه على الدعاية شهرياً فإن متوسط قيمة المبيعات المتوقعة ستتراوح بين 1.64 ، 3.76 ألف جنيه .

ب- تقدير قيمة المبيعات ص^أ عند س = 4 بفترة ثقة 95 % :

$$\text{لمر } \frac{1}{\text{س}} = \text{ص}^{\wedge} + \text{ت} (ن-2) \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$= 2.7 + 3,182 \times \sqrt{\frac{1}{5} + \frac{(\text{س} - \text{س}^{\wedge})^2}{\text{ن} - 2}} \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$= 2.7 + 3,182 \times \sqrt{\frac{1}{5} + \frac{(4 - 3)^2}{15 - 2}} \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$= 2.7 + 3,182 \times 0.7 = 2.7 + 2.22 = 4.92, 0.48$$

أى أنه عندما تتفق الشركة 4 آلاف جنيه في الشهر القادم فإن المبيعات المتوقعة في ذلك الشهر ستتراوح بين 0.48 ، 4.92 ألف جنيه .

ج- تقدير متوسط المبيعات عند س = 3 ثم عند س = 7 :

$$\text{لمر } \frac{1}{\text{س}} = \text{ص}^{\wedge} + \text{ت} (ن-2) \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$= (3 \times 0.7 + 0.1 - 3) + 3,182 \times \sqrt{\frac{1}{5} + \frac{(\text{س} - \text{س}^{\wedge})^2}{\text{ن} - 2}} \times \frac{1}{\text{ع} (\text{ص}^{\wedge})}$$

$$\begin{aligned}
 &= \sqrt{\frac{(3-3)^2}{1.0} + \frac{1}{0}} \times 0.61 \times 3.182 + 2 = \\
 &1.14 \text{ و } 2.86 = 0.86 + 2 = 0.27 \times 3.182 + 2 = \\
 &\text{وعند س. } 7 = \text{ نجد أن :} \\
 &\text{لم س.}^8 = 0.16 \times 3.182 + 4.8 = 7 = \sqrt{\frac{(3-7)^2}{1.0} + \frac{1}{0}} \times 0.16 \times 3.182 + 4.8 = \\
 &(7.4, 2.2) = 2.6 + 4.8 = 0.82 \times 3.182 + 4.8 =
 \end{aligned}$$

ملحوظة :

إذا كان الحد الأدنى لفترة الثقة سالب القيمة ، فإننا نعتبره صفراً ، لأنه من غير المنطقي أن تكون قيمة المبيعات سالبة القيمة .

تعليق :

يلاحظ أنه عندما كانت س. = -3 كان الخطأ المعياري ع(ص) = 0.27 وعندما كانت س. = 4 أصبح الخطأ المعياري 0.33 وعندما كانت س. = 7 ارتفع الخطأ المعياري إلى 0.82 أى أنه كلما ابتعدت س. عن \bar{S} كلما زاد الخطأ المعياري ، ويزداد الخطأ المعياري بصورة خطية إذا كانت س. تقع خارج مجال قيم س في العينة (لاحظ أن مجال س في العينة يتراوح بين 1 و 5) لذا ينصح أن يتم التنبؤ بقيمة ص⁸ عند قيمة س. القريبة من متوسطها \bar{S} .

مثال (١٧) شامل :

قام مدير إحدى الشركات بتحليل العلاقة بين مستوى أداء العاملين بها (ص) ومعدلهم التراكمي في الكالوريوس (س) حيث يعتقد بوجود علاقة بينهم ، فحصل على النتائج التالية :

مسلم	س	ص	س ص	س ²	ص ²
١	٣	٥	١٥	٩	٢٥
٢	٢	٤	٨	٤	١٦
٣	٤	٤	١٦	١٦	١٦
٤	١٢	٩	١٠٨	١٤٤	٨١
٥	١١	٨	٨٨	١٢١	٦٤
٦	٨	٩	٧٢	٦٤	٨١
٧	٩	٧	٦٣	٨١	٤٩
٨	٧	٨	٥٦	٤٩	٦٤
٩	٦	٥	٣٠	٣٦	٢٥
١٠	٥	٦	٣٠	٢٥	٣٦
١١	٤	٨	٣٢	١٦	٦٤
١٢	٨	٤	٣٢	٦٤	١٦
١٣	٣	٧	٢١	٩	٤٩
١٤	١٢	٦	٧٢	١٤٤	٣٦
١٥	٩	٨	٧٢	٨١	٦٤
١٦	٨	٥	٤٠	٦٤	٢٥
١٧	١١	١٠	١١٠	١٢١	١٠٠
١٨	٧	٧	٤٩	٤٩	٤٩
١٩	٨	٦	٤٨	٦٤	٣٦
٢٠	١٠	٥	٥٠	١٠٠	٢٥
ن = ٢٠	مجس = ١٤٧	مجص = ١٣١	مجس ص = ١٠١٢	مجس ² = ١٢٦١	مجص ² = ٩٢١

(١) معادلة خط الانحدار ص = أ + ب س ، حيث :

$$\begin{aligned} \text{مجص} &= \text{ن} \times \text{أ} + \text{ب} \times \text{مجس} \\ \text{مجس ص} &= \text{أ} \times \text{مجس} + \text{ب} \times \text{مجس}^2 \end{aligned}$$

وبحل تلك المعادلتين نجد أن : $\text{أ} = ٤,٥٥$ ، $\text{ب} = ٠,٢٧$

ويمكن استخدام صيغة القانون لحساب أ ، ب كما يلي :

$$ب = \frac{\text{مجم س ص} - \text{مجم س} \times \text{مجم ص} / \text{ن}}{\text{مجم س}^2 - (\text{مجم س})^2 / \text{ن}} = \frac{20/131 \times 147 - 1012}{20/1^2 (147) - 1261} = \frac{20}{20/1^2 (147) - 1261} = 0,27$$

$$أ = \text{ص} - ب \times \text{س} = 147 - \frac{131}{20} \times 0,27 = 147 - 1,79 = 145,21$$

∴ معادلة خط الانحدار ص = 0,27 + 145,21 س

(2) الخطأ المعياري للتقدير

$$\text{غير} = \sqrt{\text{مجم ص}^2 - \text{أ} \times \text{مجم ص} - \text{ب} \times \text{مجم س} \times (\text{ن} - 2)}$$

$$\frac{51,71}{18} = \sqrt{18 \div (1012 \times 0,27 - 131 \times 145,21 - 921)} = \sqrt{1,695} = 1,29$$

(3) فترة الثقة لمعامل الانحدار β عند α = 5% :

$$ب \pm ت = \beta \quad \text{ت} = (t/\alpha, \text{ن} - 2) \times \text{ع} (ب)$$

حيث ع (ب) : الخطأ المعياري للتقدير ب = غير ÷ √(مجم س - (مجم س)² / ن)

$$\text{ع} (ب) = \frac{1,695}{13,4369} = \frac{1,695}{20/1^2 (147) - 1261} = 0,126$$

$$ب \pm ت = 0,27 \pm (0,20 \times 0,18) \times 0,126$$

$$0,0347 \leq 0,00053 \leq 0,2647 \pm 0,27 = 0,126 \times 2,101 \pm 0,27 =$$

(4) اختبار β = صفر عند α = 5% باستخدام توزيع ت :

خطوات الاختبار :

١- الفرض العدمي β = صفر

٢- الفرض البديل β ≠ صفر

$$\text{ت} = \frac{\text{ب} - \beta}{\text{ع} (ب)} = \frac{0,27}{0,126} = 2,143$$

٤- ت الجدولية : ت (١٨ ، ٠.٠٢٥) = ٢,١٠١

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية .

٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .

∴ $\beta \neq 0$ ، أي أن معامل انحدار المعدل التراكمي يختلف عن الصفر .
ويمكن استخدام فترة الثقة السابقة كأسلوب آخر لاختبار $\beta = 0$ ، صفر ،
ونذلك بالبحث عن الصفر داخل حدى الثقة ، وحيث أن الصفر لا يقع
داخل حدى الثقة يتخذ قرار برفض الفرض العدمي وقبول الفرض
البديل .

(٥) اختبار $\beta = 0$ صفر باستخدام تحليل التباين :

$$\text{م.م.ك} = \text{م.ج.ص}^2 - (\text{م.ج.ص})^2 / \text{ن} = ٩٢١ - ٢٠ / (١٣١) = ٦٣$$

$$\text{م.م.} \text{ بسبب انحدار ص/س} = \text{ب(م.ج.ص ص - م.ج.ص} \times \text{م.ج.ص/ن)}$$

$$= ١٣,٢٧ - (٢٠ / ١٣١ \times ١٤٧ - ١٠١٢) = ١٣,٢٧$$

جدول تحليل التباين

المصدر	مجموع المربعات	درجات الحرية	متوسط مجموع المربعات	نسبة التباين	ف الجدولية
م.م. انحدار ص/س	١٣,٢٧	١	١٣,٢٧	٤,٨	ف(١,١٨,٥) = ٣,٠١
م.م.د (البواقي)	٤٩,٧٣	١٨	٢,٧٦٣		
م.م.ك	٦٣	١٩			

وحيث أن ف المحسوبة تتجاوز ف الجدولية عند مستوى المعنوية ٥%

، فيذا يعني رفض الفرض العدمي وقبول الفرض البديل ، أي أن β

= صفر وهو نفس القرار الذي توصلنا إليه سابقاً .

(٦) تقدير متوسط أداء العاملين بفترة ثقة ٩٥% عند :

$$(أ) س = ٤ ، (ب) س = ٧,٣٥ ، (ج) س = ١٠$$

بوضع كل قيمة من قيم س السابقة في المعادلة $ص = ٤,٥٥ + ٠,٢٧ س$

نحصل على قيمة ص^٨ المناظرة ، ، وبين الجدول التالي هذه القيم

بالإضافة إلى قيم مقادير أخرى لازمة لحل هذا المطلوب ، مع العلم بأن

$$س = ٧,٣٥ = ٢٠ \div ١٤٧$$

س	ص ^٨	ت (١٨٠,٠٢٥)	خ من	$\sqrt{\frac{١}{ن} + \frac{(س - س)^2}{محس - محس^2 / ن}}$
٤	٥,٦٣	٢,١٠١	١,٦٩٥	$٠,٣٣٤٩ = \frac{١١,٢٢}{١٨٠,٥٥} + \frac{١}{٢٠}$
٧,٣ ٥	٦,٥٣	٢,١٠١	١,٦٩٥	$٠,٢٢٣٦ = \frac{صفر}{١٨٠,٥٥} + \frac{١}{٢٠}$
١٠	٧,٢٥	٢,١٠١	١,٦٩٥	$٠,٢٩٨٢ = \frac{٧,٠٢}{١٨٠,٥٥} + \frac{١}{٢٠}$

وبوضع هذه الكميات في العلاقة التالية :

$$\mu مره = ص^٨ + ت (٧/١٨٠,٢٥ - ن) \times ع (ص^٨)$$

نصل إلى فترات الثقة الثلاث التالية :

$$١- \mu مره = ٤,٤٣٧ \& ٦,٨٢٣ \quad \text{مدى الثقة} = ٢,٣٨٦$$

$$٢- \mu مره = ٥,٧٣٨ \& ٧,٣٣١ \quad \text{مدى الثقة} = ١,٥٩٣$$

$$٣- \mu مره = ٦,١٨٨ \& ٨,٣١٢ \quad \text{مدى الثقة} = ٢,١٢٤$$

يلاحظ أن فترة الثقة تزداد اتساعاً كلما ابتعدت س. عن س ، وهذا يعني

أن تقدير المتوسط يقل الاعتماد كلما ابتعدت س. عن س ويرجع السبب في ذلك

إلى أن قيمة الخطأ المعياري للتقدير ع(ص) تصل إلى أدنى قيمة لها عندما تصبح س. = س وأن قيمته تزداد كلما ابتعدت س. عن س

(٧) تقدير قيمة ص^٨ (وليس متوسط ص^٨) بفترة ثقة ٩٥% :

عند س = ١٠ كانت ص^٨ = ٧,٢٥

المسألة: ص^٨ = ص^٨ ± ت (٠,٠٢٥, ١٨) × ع(ص^٨)

$$= 7,25 \pm 2,101 \times \sqrt{\frac{1}{180,25} + \frac{1}{20} + 1} \times \sqrt{\frac{(7,35-10)^2}{180,25}}$$

$$= 7,25 \pm 1,690 \times 1,044$$

$$= 7,25 \pm 1,766 = 3,484 \text{ ، } 10,966 \text{ ، } 3,484 \text{ مدى الثقة } = 7,432$$

يلاحظ أن فترة الثقة الأخيرة أكثر إتساعاً من أي فترة حصلنا عليها عند تقدير

متوسط ص^٨

مثال (١٨) نصف محلول

بفرض أنك حصلت على البيانات التالية :

ص	٢٠	٢٨	٤٠	٤٥	٣٧	٥٢	٥٤	٤٣	٦٥	٥٦
س	٢	٣	٥	٤	٣	٥	٧	٦	٧	٨

المطلوب :

(أ) معادلة خط انحدار ص / س

الإجابة : ص = ١٤,٢٨ + ٥,٩٤ س

(ب) تبين تقديرات معادلة خط الانحدار والخطأ المعياري لها

الإجابة : ع^٢ ص = ٤٦,٩٧ ، ع ص = ٦,٨٥

(ج) تبين التقدير (ب) والخطأ المعياري له .

الإجابة : ع^٢ (ب) = ١,٣١ ، ع (ب) = ١,١٤

(ح) اختبار β = صفر عند α = ٠,٠٥

- الاجابة β : معنوية عند $\alpha = 0.05$.
 (د) فترة الثقة للمعلمة β عند $\alpha = 0.05$.
 الاجابة $\beta = 3.31, 8.57$.
 (هـ) معامل التحديد ومعامل الارتباط
 الاجابة $r = 0.77, r^2 = 0.88$.

مثال (١٩) (نصف محلول)

لتقدير أثر برنامج تدريبي على القدرة البيعية لمندوبي المبيعات باحدى الشركات ، جمعت البيانات التالية وهي تمثل (س) : الدرجة التي حصل عليها مندوب المبيعات بعد انتهاء البرنامج التدريبي الذي استمر شهرين ، (ص) : قيمة المبيعات لهم بالآلاف جنيه خلال السنة التالية لإنهاء البرنامج التدريبي .

س	١٨	٢٦	٢٨	٣٤	٣٦	٤٢	٤٨	٥٢	٥٤	٦٠
ص	٥٤	٦٤	٥٤	٦٢	٦٨	٧٠	٧٦	٦٦	٧٦	٧٤

المطلوب :

- (أ) : معادلة خط انحدار ص / س بطريقة المربعات الصغرى

$$\text{ص} = ٤٦,٥٢٤ + ٠,٤٩٩ \text{ س}$$
- (ب) قياس الخطأ المعياري لتقدير معادلة خط الانحدار مستخدماً في ذلك كل من القيم الفعلية (ص) والقيم التقديرية (ص)

$$\text{خ}^2 = \text{مح} - (\text{ص} - \text{ص}^{\wedge})^2 \div (\text{ن} - ٢) = ١٧٠,٥٠١ \div ٨ = ٢١,٣٢$$

$$\therefore \text{خ} = ٤,٦١٦$$
- (جـ) اختبار β = صفر باستخدام تحليل التباين

$$\text{م.م.ك} = ٥٩٠,٤ \text{ ، م.م.انحدار ص/س} = ٤١٩,٧٩٦$$
 نسبة التباين (ف المسحوبة) = $١٩,٦٨٥$ (تقريباً)، ف الجدولية
 ف $(١,٨٠٥\%) = ٥,٣٢$

القرار : رفض الفرض العدمي وقبول البديل أي $\beta \neq 0$ صفر

(د) معامل التحديد : $r = 0.711 - 0.711 = 0.000$ أما معامل الارتباط $= 0.843$

(هـ) اختبار $\beta = 0$ صفر باستخدام توزيع ت

$$t = \frac{\beta - 0}{\frac{s}{\sqrt{n}}} = \frac{0.499}{\frac{21.32}{\sqrt{1683.6}}} = 0.416$$

$$t = (8, 0.025) = 2.306$$

القرار : رفض الفرض العدمي وقبول البديل

(و) تقدير β بفترة ثقة 95%

$$\beta = \bar{\beta} \pm t_{(n-1), \alpha/2} \times s_{\beta}$$

$$= 0.499 \pm 2.306 \times 0.133 = 0.238 \text{ و } 0.76$$

(ز) تقدير متوسط $\hat{\mu}$ عند $s = 39.8$

$$\bar{x} = 21.32 = \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \frac{(\sum_{i=1}^n x_i)^2}{n}$$

$$2.132 = \bar{x} - 39.8 = \bar{x} - 39.8$$

$$\bar{x} = 39.8 + 2.132 = 41.932$$

$$= 41.932 \times 2.306 + 66.4 = 100.4$$

$$= 100.4 \text{ و } 100.4$$

مثال (٣٠) (نصف محلول)

البيانات التالية تمثل نتائج دراسة في نهاية سنة ما عن تكاليف الصيانة

(ص) لألات بيع المشروبات آليا بحسب العمر الزمني (س) وقد شملت العينة

١٠ ماكينات وكانت النتائج كما يلي : س = العمر الزمني للآلة ، ص : تكاليف

الصيانة سنويا .

$$\text{محس} = 35, \text{ محس} = 1397, \text{ محس} = 5004$$

محـ س^٢ = ١٣٣ محـ ص^٢ = ١٩٦٦٦٥ ن = ١٠ =

المطلوب :

- ١- معادلة خط انحدار ص /س بطريقة المربعات الصغرى .
- ٢- قدر التكاليف المتوقعة لآلة ما بعد مرور عامين على استخدامها .
- ٣- استخدام تحليل التباين لاختبار β = صفر وحقق ما تصل إليه باستخدام اختبار ت .
- ٤- قدر β بفترة ثقة ٩٥% .

الحل:

- ١- ص = ١٠١,٦ + ١٠,٩ س
- ٢- ص^٨ عند س = ٢ هي : ص^٨ = ١٢٣,٣ جنيه
- ٣- م.م.ك = ١٥٠٤,١ ، م.م.ب = ١٢٤٨,٠٥ ، ف = ٣٨,٩٩ ، ت = ٦,٢٤
- ٤- $\beta = ١٠,٩ + ٤,٠٢٧٩ = ٦,٩ \& ١٤,٩$

تمارين

(١) أ- ارسم الخط البياني المار بالنقط التالية :

(١) (١،١) & (٥،٥) (٢) (٣،٠) ، (٠،٣)
(٣) (١،١-) & (٢،٤) (٤) (٣-، ٦-) ، (٦، ٢)

ب- حدد كل من الميل والجزء المقطوع في الخطوط السابقة بيانيا .

ج- أوجد المعادلات الرياضية التي تمر بالنقط في (١)

(٢) ارسم الخطوط المستقيمة التالية بيانيا ثم حدد الميل والجزء المقطوع

(١) ص = س + ٤ (٢) ص = -٥ - ٢س

(٣) ص = -٢س (٤) ص = ٦ + ٢س

(٣) أوجد معادلة خط انحدار ص / س ، وكذلك معادلة خط انحدار س / ص في

الحالات التالية مستخدما طريقة المربعات الصغرى :

(أ)

س	٧	٤	٦	٢	١	٣
ص	٢	٤	٢	٥	٧	٥

(ب)

س	٨	٥	٤	٦	٢	٥	٣
ص	١	٣	-٦	٣	٧	٢	٥

(ج) ارسم الشكل الانتشاري لكل من (أ) ، (ب) ثم وفق عليها

الخطوط التي توصلت إليها في (أ) ، (ب) .

(٤) بين أي المعادلات التالية تعبر عن معادلة انحدار بسيط وأيهما تعبر عن انحدار متعدد .

(أ) ص = ٣ + ٢س (ب) ص = ٣ + ٢س + ٥س

(ج) ص = ٧ - ٢س + ٣س (د) ص = -٥ - ٣س

(٥) ما هو الفرق بين الارتباط والانحدار ؟ وما هو المقصود بعبارة : انحدار ص على س ؟ حدد أيهما متغير مستقل وأيهما متغير تابع .

(٦) البيانات التالية تمثل عدد أيام الغياب (ص) ، وعدد سنوات العمل (س) في إحدى الشركات ، لعينة عشوائية من ١٠ عمال بهذه الشركة

ص	٢	٠	٥	٦	٤	٩	٢	١٥	٠	٧
س	٧	٨	٢	٣	٥	٤	٦	١٠	٤	١١

المطلوب :

- رسم الشكل الإنشائي لهذه البيانات .
- إيجاد معادلة خط انحدار ص / س بطريقة المربعات الصغرى .
- التنبؤ بعدد أيام الغياب لعامل بالشركة يعمل بها منذ أربع سنوات .
- احسب قيمة معامل التحديد ومعامل الارتباط .
- اختبر الفرض القائل بعدم وجود علاقة بين (س ، ص) عند $\alpha = 5\%$ مستخدماً أسلوب تحليل التباين

(٧) البيانات التالية تمثل عدد سنوات الخدمة (س) وعدد العاملين الذين استقالوا من العمل بإحدى الشركات (ص) :

عدد العاملين المستقبليين (ص)	٧	٦	٣	٢	٧
عدد سنوات الخدمة (س)	٤	٥	٦	٧	٣

المطلوب :

- معادلة خط الانحدار ص / س بطريقة المربعات الصغرى .
- معامل التحديد ومعامل الارتباط موضحاً الفرق بينهما ثم معامل عدم التحديد .

ج - اختبار معنوية معامل الانحدار مستخدماً في ذلك كل من أسلوب فترة الثقة وأسلوب اختبار المعنوية (توزيع ت) وذلك عند مستوى معنوية ٥% .

(٨) ترغب إحدى شركات التأمين على الحياة في تحديد العلاقة بين خبرة البائع وحجم المبيعات له . سحبت عينة عشوائية من تسعة مندوبي المبيعات حيث سجلت خبرة كل منهم بالسنوات (س) ومبيعاتهم السنوية (ص) في السنة الحالية بمئات الآلاف من الجنيهات فكانت

٩	٨	٧	٦	٥	٤	٣	٢	١	ص
٧	٥	٦	٥	٤	٣	٣	١	٢	س

المطلوب

- I- تقدير المبيعات السنوية لمندوب خبرته ١٠ سنوات .
- II- الخطأ المعياري لتقدير معادلة خط الانحدار .
- ج- اختبار $\beta = 0$ - صفر عند $\alpha = 5\%$ مستخدماً توزيع (ت) ثم دعم قرارك بأسلوب فترة الثقة .
- د - أوجد فترة الثقة ٩٥% لمتوسط المبيعات عندما تكون :
(١) س = ٤,٥ ، (٢) س = ١٠
- هـ - أوجد فترة الثقة ٩٥% للقيمة الفعلية ص عند س = ١٠ سنوات .
- و- قارن بين النتائج التي حصلت عليها في كل من (د) ، (هـ) ومعلقاً عليها .

(٩) تحتفظ إحدى الشركات بسجل عن تكلفة صيانة كل ماكينة من ماكيناتها ، بالإضافة إلى عمر كل منها . لمعرفة العلاقة بين عمر الماكينة (س)

وتكلفة الصيانة (ص) ، سحبت عينة عشوائية من ٦ ماكينات وحصلنا على البيانات التالية :

الماكينة	١	٢	٣	٤	٥	٦
س	٢	١	٣	٢	١	٣
ص	٧٠	٤٠	١٠٠	٨٠	٣٠	١٠٠

المطلوب

- I- معادلة خط انحدار ص / س بطريقة المربعات الصغرى .
- II- تحديد تكلفة الصيانة لماكينة عمرها ٤ سنوات .
- ج- مستخدما أسلوب فترة الثقة ، اختبر $\beta = \text{صفر}$ عند $\alpha = ١\%$
- د- ايجاد فترة الثقة ٩٩% لمتوسط التكلفة عند س = ٢ ثم عند س = ٤
- هـ- ايجاد فترة الثقة ٩٩% للتكلفة الفعلية ص عند س = ٢

(١٠) اجريت تجربة لتحديد العلاقة بين كمية الأمطار بالبوصة (س) ومحصول

القمح (ص) بالطن ، فحصلنا على البيانات التالية :

كمية الإمتار س	١	٢	٣	٤	٥	٥	٦	٧	٨	٩
كمية القمح ص	١	٣	٢	٥	٥	٤	٧	٦	٩	٨

المطلوب :

- أ - معادلة خط انحدار ص / س بطريقة المربعات الصغرى .
- ب- معامل التحديد ومعامل الارتباط ومعامل التحديد وملول كل منهم .
- ج- الخطأ المعياري لتقدير معادلة خط الانحدار .
- د- اختبار الفرض القائل بعدم وجود علاقة بين كمية المطر ومحصول القمح ($\alpha = ٥\%$) مستخدما في ذلك أساليب مختلفة من الاختبارات.

(١١) في عينة عشوائية حجمها ١٠ من طلبة إحدى الجامعات ، جمعت بيانات عن المعدل التراكمي في المرحلة الثانوية (س) والمعدل التراكمي في الجامعة (ص) ، فحصلنا على المعلومات التالية :

$$\begin{aligned} \text{مـ س} &= 32,2 & \text{مـ ص} &= 30 & \text{مـ س ص} &= 98,32 \\ \text{مـ س} &= 2 & \text{مـ ص} &= 105,74 & \text{مـ ص} &= 91,9 & \text{ن} &= 10 \end{aligned}$$

المطلوب :

- أ - معادلة خط انحدار ص/س .
- ب - الخطأ المعياري لتقدير معادلة خط الانحدار .
- ج - فترة الثقة ٩٥% لمعلمة الانحدار β .
- د - اختبار β = صفر مستخدماً : (١) فترة الثقة (٢) اختبار المعنوية باستخدام توزيع (ت) (٣) تحليل التباين وذلك عند $(\alpha = 5\%)$.
- هـ - قدر متوسط ص بفترة ثقة ٩٥% عند (١) س = ٣٠، (٢) س = ٤٠ .

(١٢) أجريت دراسة عن العلاقة بين كمية النتروجين في السماد (س) وكمية محصول القمح (ص) على عينة من ٢٠ فدان متماثلة ، فحصلنا على :

$$\begin{aligned} \text{مـ س} &= 230 & \text{مـ ص} &= 260 & \text{مـ س ص} &= 3490 \\ \text{مـ س} &= 2 & \text{مـ ص} &= 3144 & \text{مـ س ص} &= 3904 \end{aligned}$$

المطلوب :

- I - معادلة خط الانحدار ص/س بطريقة المربعات الصغرى
- II - الخطأ المعياري لتقدير معادلة خط الانحدار .
- ج - اختبار β = صفر مستخدماً أسلوب تحليل التباين عند $\alpha = 5\%$
- د - قدر بفترة ثقة ٩٥% متوسط كمية القمح عند س = ١١,٥ ثم عند س = ١٥ ثم حدد أي التقديرين تفضل ولماذا ؟

(١٣) (أ) إذا كانت معادلة خط انحدار ص / س هي : $ص = ٣ + ٤٢ س$ ومعادلة خط انحدار س / ص هي : $س = ٢ + ٣٠ ص$ ، أوجد ما يلي :
 ١- معامل الارتباط بين س ، ص -٢ قيمة س عندما تكون ص = ١٢
 (ب) إذا كان معامل الارتباط الخطي بين (س، ص) هو ٠,٨ ومعامل انحدار ص/س هو ١,٦ فإذا علمت أيضا أن $س = ١١$ ، $ص = ٢٠$ ، أوجد معامل انحدار س/ص وكذلك معادلتها خط انحدار ص / س ، س / ص ثم قدر قيمة ص عندما تكون س = ١٠ .

(١٤) (أ) إذا كانت معادلة خط انحدار ص / س هي : $ص = ٠,١ س + ٧,٧$ ومعادلة خط انحدار س / ص هي : $س = ٥ - ٧ ص$ أوجد س، ص
 (ب) إذا كان خطي الانحدار على الصورة التالية :

س + ٢ ص - ٥ = صفر & ٢ س + ٣ ص - ٨ = صفر
 فإذا علمت أن تباين س أي ($ع^٢ س$) = ١٢ ، أوجد :
 ١- $س$ ، $ص$ ٢- $ع$ ٣- معامل الارتباط بين س ، ص

(١٥) (أ) إذا كان معاملي خط الانحدار ص / س ، س / ص هما : ب = ٣,٢ ، د = ٠,٨ ، بين أن هناك خطأ في أحد هذه المعاملات .
 (ب) إذا كان معاملي خط الانحدار ب = ٠,٨ ، د = ٠,٢ ، فما هي قيمة معامل الارتباط

(ج) بفرض أنه في معادلة خط الانحدار البسيط ، كانت ن = ١٥ ، $ر^٢ = ٠,٨$ فما هي قيمة وسيلة الاختبار ف إذا ما استخدم أسلوب تحليل التباين ، وما هي نتيجة اختبار β -صفر عند $\alpha = ٥\%$.

(د) بفرض أن معامل التحديد $R^2 = 0.8$ وأن الانحراف المعياري للظاهرة (ص) هو ٦ ، فما هو قيمة الخطأ المعياري لتقدير معادلة خط الانحدار ، غير ؟

(١٦) (أ) استخدم البيانات التالية في إيجاد معادلتى خط انحدار ص/س ، س/ص

ثم اعرضهما بيانيا موضحا منلول نقطة التقاطع بينهما

محـس = ٥٨٠ ، محـص = ٣٧٠ ، محـس ص = ١١٤٩٤

محـس^٢ = ٤١٦٥٨ ، محـص^٢ = ١٧٢٠٦ ، ن = ١٢

(ب) في دراسة على عينة من ١٠ مفردات ، حصلنا على النتائج التالية:

ص = ٨ + ٠.٤ س ، خ^٢ مر = ١.٥ ، ع^٢ مر = ٤

المطلوب :

١- ميل خط انحدار س/ص ٢- معامل الارتباط بين (س ، ص)

(١٧) من المعتاد في مجال التجارة إعطاء خصم عند شراء كمية كبيرة من السلعة . البيانات التالية توضح هذه الحقيقة .

٤	٦	٧	٨	١٠	سعر الوحدة (ص)
٥	٤	٣	٢	١	عدد الوحدات المشتراة (س)

المطلوب :

أ- عرض هذه البيانات بيانيا موضحا الجزء المقطوع وميل خط

الانحدار بعد أن تكون قد وفقت لها معادلة خط انحدار ص/س

بطريقة المربعات الصغرى .

ب- اختبار β = صفر عند $\alpha = 5\%$ مستخدما اختبار ت ثم أعد

الاختبار مرة أخرى باستخدام تحليل التباين . هل هناك اختلاف في

القرار الذي تتخذه ؟ ولماذا ؟

جـ - قدر متوسط ص بفترة ثقة ٩٥% عند س = ٣ ثم عند س = ٥
ثم حدد أي الفترتين تفضل ولماذا ؟

(١٨) الجدول التالي يبين كمية المعروض من سلعة ما (ص) عند اسعار مختلفة (س) :

ص	١٢	١٤	١٠	١٣	١٧	١٢	١١	١٥
س	٥	١١	٧	٨	١١	٧	٦	٩

المطلوب :

- أ- معادلة خط الانحدار ص/س
ب- اختبر المعنوية الاحصائية لميل خط الانحدار β باستخدام توزيع ت
ثم باستخدام فترة الثقة ، ثم دعم ذلك مستخدماً أسلوب تحليل التباين.
جـ - أوجد معامل التحديد ومعامل عدم التحديد ومدلول كل منهم .
د- قدر قيمة ص^٨ بفترة ثقة ٩٥% عند س = ١٠

ملحوظة :

عليك التأكد من الاجابات التالية :

أ- ص = $٦,٤٤ + ٠,٨٢ س$

ب- ت = ب ÷ ع (ب) = $٠,٨٢ ÷ ٠,٢٥ = ٣,٢٨$

جـ - ر = $٠,٦٤٠٥$

د- ص^٨ = $١٠ \times ٠,٨٢ + ٦,٤٤ = ١٤,٦٤$

م.م. = $١٤,٦٤ + ت \times (٠,٢٥, ١٠) \times ع (ص) = ١٢,٩٤١٦$

حيث ع^٨ (ص) = $\left(\frac{(٨-١٠)}{٣٤} + \frac{١}{٨} + ١ \right) \frac{١٢,٩٤١٦}{٦} = ٢,٦٧$

م.م. = $١٨,٦٣ = ١٠,٦٥ + ١,٦٣ \times ٢,٤٥ + ١٤,٦٤$

(١٩) الجدول التالي يوضح توزيع عينة من الأسر حسب حجم الدخل السنوى بالآلف جنيه (س) وحجم الإنفاق بالآلف جنيه (ص) والمطلوب :

١- معادلة خط انحدار ص/س .

٢- تقدير الإنفاق المتوقع لأسرة يبلغ دخلها السنوى ٦ (آلف جنيه).

ص س	-٣	-٥	-٧	-٩	المجموع
-٨	٣	٤	٢		٩
-١٠	٤	٦	٢		١٢
-١٢		٥	٣	٦	١٤
-١٤			٤	٣	٧
١٨-١٦			٥	٣	٨
المجموع	٧	١٥	١٦	١٢	٥٠

الفصل الرابع

تحليل الانحدار المتعدد

Multiple Regression Analysis

- (١) مقدمة
- (٢) فروض نموذج الانحدار .
- (٣) تقدير معاملات الانحدار الجزئية .
- (٤) معامل التحديد ومعامل التحديد المعدل .
- (٥) الخطأ المعياري للتقدير .
- (٦) معاملات الارتباط الجزئية .
- (٧) الاستدلال الإحصائي عن معالم خط الانحدار .
- (٨) اختبار المعنوية الكلية للانحدار : مدخل تحليل التباين .
- (٩) اختبار مساهمة المتغيرات التفسيرية: مبدأ مجموع المربعات الإضافي .
- (١٠) بعض مشاكل استخدام تحليل الانحدار :
 - ١- مشكلة عدم ثبات التباين .
 - ٢- مشكلة الازدواج الخطي .
 - ٣- مشكلة الارتباط الذاتي للبيانات .

تمارين

الفصل الرابع

تحليل الانحدار المتعدد

Multiple Regression Analysis

(1) مقدمة Introduction

في الانحدار الخطى البسيط كنا نهتم بدراسة تأثير متغير مستقل واحد S على المتغير التابع Y ، لكن من الناحية الواقعية والعملية فإن المتغير التابع دائماً يكون واقعاً تحت تأثير عدة متغيرات مستقلة أو تفسيرية وليس متغير مستقل واحد. فمثلاً عند دراسة العلاقة بين الدخل والإنفاق، كنا نفترض ضمناً أن الدخل (S) فقط يؤثر في الإنفاق (Y)، لكن النظرية الاقتصادية لهذه العلاقة نادراً ما تكون بهذه البساطة، فبجانب الدخل، هناك عدداً من المتغيرات التفسيرية الأخرى تؤثر أيضاً في الإنفاق مثل: ثروة المستهلك، المستوى التعليمي، المستوى الاجتماعي... إلخ. مثال آخر: الطلب على سلعة ما (Y) لا يعتمد فقط على سعرها (S) ولكن يعتمد أيضاً على أسعار السلع الأخرى سواء المتنافسة أو المتكاملة معها، ويعتمد أيضاً على دخل المستهلك وعلى الحالة الاجتماعية... إلخ، إلى غير ذلك من الأمثلة، لذا فإننا نحتاج إلى تطوير وتوسيع نموذج الانحدار البسيط ليغطي عدداً أكبر من المتغيرات المستقلة، وهذا يقودنا إلى مناقشة نموذج الانحدار المتعدد، أي النموذج الذي يضم العديد من المتغيرات المستقلة والتي تؤثر في متغير تابع واحد (Y). وسوف نكتفي بأبسط نموذج لانحدار متعدد وهو النموذج الذي يضم متغير تابع واحد (Y) ومتغيرين مستقلين S_1 ، S_2 ، وصورته العامة في مجتمع الدراسة هي:

$$Y = \beta_0 + \beta_1 S_1 + \beta_2 S_2 + \psi \quad (1)$$

حيث:

β_0 : مقدار ثابت أو هو قيمة Y عندما تكون $S_1 = S_2 = 0$ - صفر

β_1 : معامل انحدار جزئي للمتغير ص على س₁ مع اعتبار س₂ متغير ثابت أو تعرف β_2 بصورة أخرى على أنها مقدار التغير الذي يحدث في التابع ص عندما يتغير المتغير المستقل س₁ بوحدة واحدة مع فرض ثبات تأثير المتغير المستقل الآخر س₂.

β_2 : معامل انحدار جزئي للمتغير ص على س₂ بفرض ثبات س₁ أو هو مقدار التغير في ص عندما تتغير س₂ بوحدة واحدة على فرض ثبات تأثير س₁.
 ψ : الحد العشوائي أو البواقي .

وعند التمثيل البياني لمعادلة خط الانحدار المتعدد على ثلاث محاور مستعمدة فإننا نحصل على ما يسمى بمستوى الانحدار Regression plane والجزء المشترك المقطوع من المحاور الثلاث المتعامدة عبارة عن المقدار الثابت β_0 .

(٣) فروض نموذج الانحدار : Assumptions of the Model

لا تختلف هذه الفروض كثيراً عن فروض نموذج الانحدار البسيط ويمكن تلخيصها على النحو التالي :

- ١- توقع (ψ) = صفر .
- ٢- تباين (ψ) = σ^2 = مقدراً ثابت .
- ٣- تغاير الحدود العشوائية $(\psi_i \text{ و } \psi_j)$ = صفر .
- ٤- تغاير $(\psi, \text{س}_1)$ = تغاير $(\psi, \text{س}_2)$ = صفر
- ٥- لا يوجد ارتباط ذاتي بين المتغيرات المستقلة أي تغاير $(\text{س}_1, \text{س}_2)$ = صفر
- ٦- التوزيع الاحتمالي للمتغير التابع ص هو التوزيع الطبيعي .

(٣) تقدير معاملات الانحدار الجزئية بطريقة المربعات الصغرى :

OLS Estimators

لتقدير المعالم المجهولة في نموذج الانحدار المتعدد $\beta_0, \beta_1, \beta_2$ فإننا نستخدم طريقة المربعات الصغرى والتي تعرضنا لها في الباب السابق . وفق

هذه الطريقة نقوم بسحب عينة عشوائية من مجتمع الدراسة ويسجل لكل مفردة فيها ثلاث قراءات تتعلق بالمتغيرات الثلاث ص_١، ص_٢، و يكون نموذج الانحدار بدلالة رموز العينة على النحو التالي :

$$\text{ص} = \text{ب} + \text{ب}_١\text{س}_١ + \text{ب}_٢\text{س}_٢ + \text{خ} \quad (٢) \dots$$

حيث خ هو حد البواقي أى العوامل العشوائية . وكما بينا من قبل فإن طريقة المربعات الصغرى تبحث فى تدينه مجموع مربعات البواقي ، مج - خ^٢ إلى أقل قيمة ممكنة أى :

مج - خ^٢ = مج (ص - ب - ب_١س_١ - ب_٢س_٢)^٢ أقل ما يمكن وعن طريق مفاضلة مج - خ^٢ جزئياً بالنسبة إلى ب ثم بالنسبة إلى ب_١، ثم بالنسبة إلى ب_٢، ومساواة الناتج بالصفر ، نحصل على ثلاث معادلات وتسمى بالمعادلات الطبيعية وهذه المعادلات هى :

$$\begin{cases} \text{مج ص} = \text{ن ب} + \text{ب}_١\text{مج س}_١ + \text{ب}_٢\text{مج س}_٢ \\ \text{مج ص س}_١ = \text{ب مج س}_١ + \text{ب}_١\text{مج س}_١^٢ + \text{ب}_٢\text{مج س}_١\text{س}_٢ \\ \text{مج ص س}_٢ = \text{ب مج س}_٢ + \text{ب}_١\text{مج س}_١\text{س}_٢ + \text{ب}_٢\text{مج س}_٢^٢ \end{cases} \quad (٣) \dots$$

وبحل تلك المعادلات آنياً بأى طريقة جبرية ، نحصل على تقديرات المربعات الصغرى ب ، ب_١ ، ب_٢ . جدير بالذكر ان هناك طريقة تقليدية للوصول إلى تلك المعادلات الثلاث تتم على ثلاث مراحل : الأولى جمع المعادلة (٢) ن من المرات والثانية ضرب طرفى المعادلة (٢) فى المتغير المستقل س_١، ثم جمعها والمرحلة الثالثة والأخيرة هى ضرب طرفى المعادلة الأصلية (٢) فى المتغير المستقل س_٢، ثم جمع الطرفين بذلك نصل إلى صيغ المعادلات الثلاث الطبيعية السابقة .

طريقة أخرى مبسطة لحساب معاملات الانحدار الجزئية :

إذا كانت البيانات المتاحة عن المتغيرات ص_١، ص_٢ ذات أعداد كبيرة ويصعب التعامل معها مباشرة يدوياً (كتربيع هذه الأعداد أو ضربها مئتي

مثلى) فإننا نلجأ إلى طريقة الانحرافات حيث تختزل تلك المتغيرات عن طريق طرح الأوساط الحسابية منها أى :

$$\bar{C} = \bar{C} - \bar{C}, \quad \bar{C}_1 = \bar{C}_1 - \bar{C}, \quad \bar{C}_2 = \bar{C}_2 - \bar{C}$$

وعلى ذلك تختزل معادلة الانحدار المتعدد إلى الصور التالية :

$$\bar{C} = \bar{C}_1 + \bar{C}_2 + \bar{C}_3 \quad (4) \dots$$

وبالتالى يمكن الحصول على الثوابت \bar{C}_1, \bar{C}_2 من خلال حل المعادلتين :

$$\begin{cases} \bar{C} - \bar{C}_1 = \bar{C}_2 + \bar{C}_3 \\ \bar{C} - \bar{C}_2 = \bar{C}_1 + \bar{C}_3 \end{cases} \quad (5) \dots$$

أما الثابت \bar{C} فتحصل عليه من المعادلة التالية :

$$\bar{C} = \bar{C}_1 + \bar{C}_2 + \bar{C}_3 \quad (6) \dots$$

مثال (1) :

استخدم البيانات الفرضية التالية فى إيجاد معادلة خط انحدار $\bar{C}/\bar{C}_1, \bar{C}_2$ ثم قدر قيمة المتغير التابع \bar{C} عندما تكون $\bar{C}_1 = 7, \bar{C}_2 = 10$

\bar{C}_1	\bar{C}_2	\bar{C}
3	3	5
1	2	4
8	4	6

الحل :

\bar{C}_1	\bar{C}_2	\bar{C}	$\bar{C}_1 - \bar{C}_1$	$\bar{C}_2 - \bar{C}_2$	$\bar{C} - \bar{C}$	\bar{C}_1^2	\bar{C}_2^2	$\bar{C}_1 \bar{C}_2$	$\bar{C}_1 \bar{C}$	$\bar{C}_2 \bar{C}$
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4	-2	-1	-1	1	4	-2	4	8
8	4	6	5	1	1	64	16	20	48	24
3	3	5	0	0	0	9	9	0	15	15
1	2	4</								

بالتعويض بهذه المجاميع في المعادلات الثلاث الطبيعية رقم (٣) :

$$20 = 5ب + 15ب_1 + 25ب_2$$

$$76 = 15ب + 55ب_1 + 81ب_2$$

$$109 = 25ب + 81ب_1 + 55ب_2$$

وبحل تلك المعادلات بأى طريقة جبرية كالحذف أو المحددات أو المصفوفات ،

نحصل على التقديرات ب ، ب₁ ، ب₂ على النحو التالى :

$$ب = 4 ، ب_1 = 2,5 ، ب_2 = -1,5$$

وعلى ذلك تصبح معادلة خط الانحدار على الصورة :

$$\text{ص} = 4 + 2,5ب_1 - 1,5ب_2$$

ولتقدير قيمة ص عند ب₁ = 7 ، ب₂ = 10

$$\text{ص} = 4 + 2,5 \times 7 - 1,5 \times 10 = 6,5$$

(٤) معامل التحديد ومعامل التحديد المعدل R² and Adjusted R²

معامل التحديد Coefficient of Determination هو ذلك المعامل

الذي يبين نسبة التغير في المتغير التابع والتي يمكن تفسيرها بسبب وجود

المتغيرات المستقلة أو بمعنى آخر معامل التحديد يعطى نسبة تأثير المتغيرات

المستقلة على المتغير التابع .

$$R^2 = \frac{\text{التغير المفسر بسبب وجود المتغيرات المستقلة}}{\text{التغير الكلى فى التابع ص}} = \frac{\text{مجم}(\text{ص} - \text{ص}^{\wedge})^2}{\text{مجم}(\text{ص} - \text{ص}^{\text{مجم}})^2} \dots (٧)$$

حيث ص : متوسط قيم ص ، ص[∧] : القيم التقديرية الناتجة من معادلة خط

الانحدار المتعدد عند القيم المعطاة لـ ب₁ ، ب₂ وهناك صيغة مبسطة لحساب R²

$$R^2 = \frac{\text{ب} \cdot \text{مجم}(\text{ص} - \text{ص}^{\wedge}) + (\text{ب}_1 - 1) \cdot \text{ب} + \text{ب}_2 \cdot \text{مجم}(\text{ص} - \text{ص}^{\wedge})}{\text{مجم}(\text{ص} - \text{ص}^{\text{مجم}})^2} \dots (٨)$$

$$= \frac{\text{ب} \cdot (\text{مجم} \text{ص} - 1 \cdot \text{مجم} \text{ص} / \text{ن}) + \text{ب}_1 \cdot (\text{مجم} \text{ص} - 2 \cdot \text{مجم} \text{ص} / \text{ن}) + \text{ب}_2 \cdot (\text{مجم} \text{ص} - 3 \cdot \text{مجم} \text{ص} / \text{ن})}{\text{مجم} \text{ص} - 4 \cdot (\text{مجم} \text{ص})^2 / \text{ن}}$$

... (٩)

وإذا اخترنا المتغيرات بأوساطها الحسابية نصل إلى :

$$R^2 = \frac{\text{مج ح}^2 + \text{ب}^2 \text{مج ح}^2}{\text{مج ح}^2} \dots (10)$$

وتطبيقاً لذلك على المثال السابق وبواسطة المعادلة (٩) :

$$R^2 = \frac{2,5 - (76 - 5/10 \times 20 - 1,5 - 1,09) (5/20 \times 20 - 1,09)}{5/20 - 1,08}$$

$$R^2 = \frac{2,5 - (16 - 1,5 - 1,09) (1,5 - 1,09)}{2,5 - 1,08} = \frac{13,5 - 4,0}{2,5 - 1,08} = \frac{9,5}{1,42} = 0,946$$

أي أن ٩٤,٦ % من التغيرات الكلية في التابع ص ترجع إلى وجود المتغيرات المستقلة س_١ ، س_٢ أو بمعنى آخر المتغيرات المستقلة س_١ ، س_٢ تؤثر بنسبة ٩٤,٦ % في المتغير التابع أما الباقي ٥,٤ % فترجع إلى تأثير العوامل العشوائية أي تأثير البواقي . أما الجذر التربيعي لمعامل التحديد فيعطي معامل الارتباط المتعدد ، أي أن $R = \sqrt{0,946} = 0,973$

يلاحظ أنه إذا كان هناك متغير مستقل ثالث س_٣ ، فإننا نضيف حد آخر في بسط أي معادلة من معادلات معامل التحديد (٨ ، ٩ ، ١٠) يمثل أثر هذا المتغير على العلاقة الكلية، وبالتالي تزداد قيمة معامل التحديد ، لأن البسط يزداد ويظل المقام ثابتاً لا يتغير ، فمثلاً إذا كان في النموذج متغير مستقل س_٣ فإن R^٢ تكون على الصورة التالية مثلاً :

$$R^2 = \frac{\text{مج (ص - ص)} (س_١ - س_١) + \text{ب}^2 \text{مج (ص - ص)} (س_٢ - س_٢) + \text{ب}^2 \text{مج (ص - ص)} (س_٣ - س_٣)}{\text{مج (ص - ص)} (س_١ - س_١) + \text{ب}^2 \text{مج (ص - ص)} (س_٢ - س_٢) + \text{ب}^2 \text{مج (ص - ص)} (س_٣ - س_٣)}$$

معنى هذا أن معامل التحديد يتأثر بعدد المتغيرات المستقلة الداخلة في تركيب نموذج الانحدار ، ويعد ذلك عيباً أو نقصاً في قيمة معامل التحديد ، لأن

بعض هذه المتغيرات المستقلة قد يكون غير معنوي التأثير ، ولتلافى هذا القصور أو النقص يتعين علينا تصحيح أو تعديل معامل التحديد R^2 Adjusted وذلك بأن نأخذ في الحسبان النقص الناشئ في درجات الحرية عند إضافة متغيرات مستقلة جديدة .

وللتوضيح : في الانحدار البسيط والذي يشمل متغير تابع ص ومتغير مستقل واحد س ، كانت درجات حرية البواقي (الخطأ العشوائي) هي (ن-٢) أما في الانحدار المتعدد والذي يشمل متغير تابع ص ومتغيرين مستقلين (س١ ، س٢) فإن درجات حرية البواقي سنجدها (ن-٣) أي أنه بزيادة المتغيرات المستقلة تنخفض درجات حرية البواقي وعليه فمن المتوقع أن تنخفض قيمة معامل التحديد المعدل (يرمز له بالرمز R^2) عن قيمته قبل التعديل :

$$R^2 = 1 - \frac{(1 - R^2) \times (n - 1)}{n - k} \quad (١١) \dots$$

حيث :

ن : حجم العينة ، ك : عدد معالم نموذج الانحدار بما فيها المقدار الثابت

R^2 : معامل التحديد

وبلاحظ على المعادلة (١١) ما يلي :

- ١- عندما تكون ك = ١ فإن $R^2 = R^2$.
 - ٢- عندما تكون ك < ١ فإن $R^2 > R^2$.
 - ٣- عندما تكون ن كبيرة جداً ومع ثبات قيمة ك فإن $R^2 \approx R^2$.
 - ٤- عندما تكون ن صغيرة جداً وفي نفس الوقت ك أكبر من ن فإن $R^2 > R^2$.
- وقد تكون سالبة القيمة أي قد يكون معامل التحديد المعدل سالب القيمة (لاحظ أن معامل التحديد R^2 لا يمكن أن يكون سالب القيمة إذ أن حدة الأدنى هو الصفر) . في هذه الحالة فإننا نعتبر معامل التحديد المعدل

مساوياً للصفر . (تذكر أن معامل التحديد R^2 تتراوح قيمته بين الصفر والواحد الصحيح)
وبحساب معامل التحديد المعدل للمثال السابق نجد أن : $R^2 = 0.946$ ، $n = 5$ ،
ك = 3 (وهي ب ، ب ، ب ، ب ، ب) . باستخدام المعادلة (١١) نصل إلى معامل التحديد المعدل R^2 :

$$\frac{4}{5} \times 0.946 - 1 = \frac{1-5}{3-5} \times (0.946 - 1) - 1 = R^2$$

$$0.892 = 0.108 - 1 =$$

ومن الواضح أن R^2 أقل من R^2

(٥) الخطأ المعياري للتقدير Standard Error of the Regression Model

إذا كانت \hat{y} تمثل التقديرات المقابلة للقيم الفعلية y ، والنتيجة عن استخدام معادلة خط الانحدار المتعدد عند القيم المعطاة للمتغيرات المستقلة (س ، س ، س) فإن الفرق بين القيم الفعلية y والقيم التقديرية \hat{y} ، يمثل الخطأ المعياري لتقدير نموذج الانحدار ويقاس على النحو التالي:

$$\text{خطأ} = \sqrt{\frac{\text{مجموع (ص - ص) }^2}{n - 3}} \quad (١٢) \dots$$

ويسمى المقام (ن - 3) بـ درجات الحرية ، وهو عبارة عن حجم العينة n مطروحاً منها عدد المعالم المطلوب تقديرها من بيانات العينة وهم ثلاثة : (ب ، ب ، ب) . وقد يستخدم البعض n بدلاً من (ن - 3) بدافع السهولة ، لكن الدقة تستوجب استخدام (ن - 3) . وتقدير الخطأ المعياري وفق العلاقة (١٢) قد يمثل صعوبة حسابية خاصة وأن القيم التقديرية \hat{y} غالباً ما تكون قيم كسرية ، لذا فهناك صيغة أخرى بديلة أكثر سهولة في الاستخدام (تذكر أن مربع الخطأ المعياري للتقدير يسمى بتباين البواقي) هي :

$$\text{غير} = \sqrt{\frac{\text{مجمد من } 1 \text{ بـ } 1 \text{ مجمد من } 1 \text{ بـ } 1 \text{ مجمد من } 1 \text{ بـ } 1}{(3-1)}} = \sqrt{\frac{1}{2}} = 0,7071 \dots$$

وتطبيقاً لذلك على بيانات المثال السابق نجد أن :

$$\begin{aligned} \text{غير} &= \sqrt{\frac{(3-1) \div (10,9 \times (1,5-1) - 76 \times 2,5 - 20 \times 4 - 10,8)}{2 \div (270 - 271,5)}} = \sqrt{\frac{1,5}{2}} = 0,8660 \dots \\ &= 0,8660 \dots \end{aligned}$$

تعليل :

يلاحظ أن قيمة الخطأ المعياري هنا لا بد وأن تكون أقل من الخطأ المعياري ، لو كنا لكتفينا بمتغير مستقل واحد ، أي عند استخدام الانحدار البسيط ، ويرجع هذا النقص في الخطأ المعياري للتقدير ، إلى إضافة متغير مستقل جديد إلى نموذج الانحدار البسيط . وعلى ذلك فالخطأ المعياري للتقدير في نموذج الانحدار المتعدد هو أقل دائماً من الخطأ المعياري في نموذج الانحدار البسيط.

وكما بينا من قبل في حالة الانحدار البسيط ، العلاقة بين الخطأ المعياري للتقدير حول خط الانحدار ومعامل الارتباط البسيط ، فهناك أيضاً علاقة مماثلة بين الخطأ المعياري حول مستوى الانحدار ومعامل الارتباط المتعدد على الصورة :

(١٤)...

$$\text{غير} = \frac{1-ن}{3-ن} \times \text{غير} \times (1-ر^2 \text{ من } (٣))$$

ومن ناحية أخرى :

(١٥)...

$$\text{غير} = \frac{1-ن}{3-ن} \times \frac{\text{غير}}{\text{غير}} \times (1-ر^2 \text{ من } (٣))$$

حيث :

خ^٢ م : مربع الخطأ المعياري للتقدير أو تباين البواقي Residuals .

ع^٢ م : تباين مفردات المتغير التابع ص حول متوسطها ص .

ر^٢ م (٣١) : مربع معامل الارتباط المتعدد بين ص وكل من س_١ ، س_٢ أى معامل التحديد المتعدد (الأرقام ١ ، ٢ تعنى س_١ ، س_٢) . والعلاقات (١٤) ، (١٥) هى علاقات إضافية لحساب الخطأ المعياري ومعامل الارتباط المتعدد كل بدلالة الآخر ، ويلاحظ أنه إذا كان حجم العينة ن أكبر من ٣٠ فمن الممكن إهمال المعامل $\frac{1-n}{3-n}$

وتطبيقاً على ذلك ، وجدنا أن الخطأ المعياري للتقدير فى المثال السابق كان :
ح^٢ م = ٠.٧٥ ، فما هى قيمة معامل الارتباط المتعدد ؟ لتطبيق المعادلة (١٥) ، يقتضى الأمر إيجاد تباين ص أى ع^٢ م حيث :

ع^٢ م = مج(ص - ص̄) ÷ (ن-١) = [مج(ص - ص̄) ÷ (ن-١)] ÷ (١-٠)
ومن بيانات المثال (١) :

ع^٢ م = [١٠.٨ - (٢٠) × ٥] ÷ ٤ = ٢٨ - ٤ = ٧ ، ومن المعادلة (١٥) :

$$ر^2 م (٣١) = ١ - \frac{٣-٥}{١-٥} \times \frac{٠.٧٥}{٧} = ١ - ٠.٠٥٤ = ٠.٩٤٦$$

أما معامل الارتباط المتعدد $\sqrt{٠.٩٤٦} = ٠.٩٧٣$ وهذه النتائج سبق الحصول عليها عند الحديث عن معامل التحديد . من ناحية أخرى إذا فرض و علمنا معامل الارتباط المتعدد ، فإنه يمكن استخدام المعادلة (١٤) في إيجاد الخطأ المعياري للتقدير .

(٦) معاملات الارتباط الجزئية : Partial Correlation Coefficients

يقيس معامل الارتباط الجزئي قوة العلاقة الارتباطية بين المتغير التابع ص و أحد المتغيرات المستقلة بعد حذف تأثير - أو مع تثبيت - المتغيرات المستقلة الأخرى . فمثلاً ر_{٢٠١} تعبر عن الارتباط الجزئي بين المتغير التابع

ص و المتغير المستقل س_١ مع ثبات تأثير المتغير المستقل الثاني س_٢ و يعرف ر_{٢٠٢} على نفس النمط . وتتراوح قيمة معامل الارتباط الجزئي بين -١ ، +١ (كما هو الحال بالنسبة لمعاملات الارتباط البسيطة) . وتقاس معاملات الارتباط الجزئية إما بدلالة معامل الارتباط المتعدد أو بدلالة معاملات الارتباط الخطية البسيطة على النحو التالي :

$$ر_{٢٠١} = \frac{\sqrt{\frac{١ - ر_{٢١}^٢}{١ - ر_{٢٢}^٢}}}{١} \quad (١٦)...$$

$$ر_{١٠٢} = \frac{\sqrt{\frac{١ - ر_{٢١}^٢}{١ - ر_{١١}^٢}}}{١} \quad (١٧)...$$

حيث ر_{٢١} : مربع معامل الارتباط المتعدد
ر_{٢٢} ، ر_{١٢} مربع معامل الارتباط البسيط بين ص ، س_١ وكذلك بين ص ، س_٢ . (الأرقام ١ ، ٢ تعني المتغيرات س_١ ، س_٢) فمثلا
مجم ص س_١ - مجم ص س_٢ × مجم س_١ س_٢ / ن
$$ر_{٢٠١} = \frac{\sqrt{\text{مجم ص س}_1 - \frac{(\text{مجم ص})^2}{ن}}}{\sqrt{\text{مجم ص س}_2 - \frac{(\text{مجم ص})^2}{ن}}}$$
 وبنفس الطريقة يمكن إيجاد المعاملات ر_{٢٠٢} ، ر_{١٠٢} .
من ناحية أخرى ، يمكن قياس الارتباط الجزئي بدلالة معاملات الارتباط

البسيطة على النحو التالي :

$$ر_{٢٠١} = \frac{ر_{٢١} - ر_{٢٢} ر_{١٢}}{\sqrt{١ - ر_{٢٢}^٢} \sqrt{١ - ر_{١١}^٢}} \quad (١٨)...$$

$$ر_{١٠٢} = \frac{ر_{١٢} - ر_{١١} ر_{٢٢}}{\sqrt{١ - ر_{١١}^٢} \sqrt{١ - ر_{٢٢}^٢}} \quad (١٩)...$$

بالطبع لا بد وأن تختلف قيمة معامل الارتباط الجزئي عن قيمة معامل الارتباط البسيط ، وذلك بسبب إدخال متغيرات مستقلة جديدة في نموذج الانحدار . جدير

بالذكر أن معاملات الارتباط الجزئية تستخدم في تحديد الأهمية النسبية للمتغيرات المفسرة (المستقلة) المختلفة في نموذج الانحدار المتعدد ، من ناحية أخرى ، يمكن استنتاج علاقة أخرى إضافية لمعامل الارتباط المتعدد بدلالة معاملات الارتباط الخطية البسيطة وهي على الصورة .

$$r_{(1)} = \frac{[r_{12}^2 + r_{13}^2 - r_{23}^2 - r_{12}r_{13}r_{23}]}{[1 - r_{23}^2]} \quad (20) \dots$$

مثال (٣)

أولاً : بفرض أنك حصلت على البيانات التالية :

$$r_{12} = 0.7, r_{13} = 0.8, r_{23} = 0.9$$

أوجد معاملات الارتباط الجزئية والكلية

ثانياً : بفرض أنك حصلت على البيانات التالية :

$$r_{12} = 0.9, r_{13} = 0.7, r_{23} = 0.6$$

أوجد معاملات الارتباط الجزئية .

الحل :

$$\text{أولاً : } r_{12} = 0.7, r_{13} = 0.8, r_{23} = 0.9$$

بالتعويض في المعادلات (١٨) ، (١٩)

$$r_{12} = \frac{0.9 \times 0.8 - 0.7}{\sqrt{0.81 - 1} \sqrt{0.64 - 1}} = \frac{0.72 - 0.7}{\sqrt{0.81 - 1} \sqrt{0.64 - 1}} = 0.1$$

$$r_{13} = \frac{0.9 \times 0.7 - 0.8}{\sqrt{0.81 - 1} \sqrt{0.49 - 1}} = \frac{0.63 - 0.8}{\sqrt{0.81 - 1} \sqrt{0.49 - 1}} = -0.076$$

$$r_{23} = \frac{0.9 \times 0.7 - 0.8}{\sqrt{0.81 - 1} \sqrt{0.49 - 1}} = \frac{0.63 - 0.8}{\sqrt{0.81 - 1} \sqrt{0.49 - 1}} = -0.076$$

أما معامل الارتباط المتعدد فهو على الصورة رقم (٢٠)

$$R_{(٢١)} = \sqrt{\frac{[R_{٢١}^2 - 1] \div [R_{٢١} \times R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٢}^2 + R_{٢٣}^2]}{[R_{٢١}^2 - 1] \div [R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٣}^2 + R_{٢١}^2]}} = R_{(٢٢)}$$

$$R_{(٢٢)} = \sqrt{\frac{[R_{٢٢}^2 - 1] \div [R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٣}^2 + R_{٢١}^2]}{[R_{٢٢}^2 - 1] \div [R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٣}^2 + R_{٢١}^2]}} = R_{(٢٣)}$$

$$R_{(٢٣)} = \sqrt{\frac{[R_{٢٣}^2 - 1] \div [R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٣}^2 + R_{٢١}^2]}{[R_{٢٣}^2 - 1] \div [R_{٢٢} \times R_{٢٣} \times 2 - R_{٢٣}^2 + R_{٢١}^2]}} = R_{(٢٤)}$$

ثانياً : $R_{(٢١)} = ٠,٩$ ، $R_{(٢٢)} = ٠,٧$ ، $R_{(٢٣)} = ٠,٦$

بالتعويض في المعادلات (١٦) ، (١٧) :

$$R_{(٢٠)} = \sqrt{\frac{[R_{(٢١)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}{[R_{(٢١)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}} = \sqrt{\frac{[0,9^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}{[0,9^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}} = ٠,٦٤$$

$$R_{(٢١)} = \sqrt{\frac{[R_{(٢١)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}{[R_{(٢١)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}} = \sqrt{\frac{[0,9^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}{[0,9^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}} = ٠,٨٣٨$$

$$R_{(٢٢)} = \sqrt{\frac{[R_{(٢٢)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}{[R_{(٢٢)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}} = \sqrt{\frac{[0,7^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}{[0,7^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}} = ٠,٧٩٢$$

$$R_{(٢٣)} = \sqrt{\frac{[R_{(٢٣)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}{[R_{(٢٣)}^2 - 1] \div [R_{(٢١)} \times R_{(٢٢)} \times 2 - R_{(٢٢)}^2 + R_{(٢٣)}^2]}} = \sqrt{\frac{[0,6^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}{[0,6^2 - 1] \div [0,9 \times 0,7 \times 2 - 0,7^2 + 0,6^2]}} = ٠,٧٩٢$$

(٧) الاستدلال الإحصائي عن معالم خط الانحدار

Inferences About The β Parameters

بيننا من قبل أن β تقيس مقدار التغير في التابع ص عندما تتغير س١ بوحدة واحدة مع ثبات تأثير المتغير المستقل الآخر س٢ ، كذلك تبين β مقدار التغير في التابع ص الناشئ عن التغير في س٢ بوحدة واحدة مع ثبات س١ . لكن إلى أي مدى يمكن الاعتماد على تقديرات تلك المعالم أي على ب١ ، ب٢ في التنبؤ بقيمة المتغير التابع ص؟ الإجابة تكمن في عملية الاستدلال الإحصائي حول تلك المعالم أي حول β ، β ، β ، (الاستدلال عن المعلمة β - المقدار الثابت -

ليست موضع اهتمام ولا يدخل تقديرها في أى عمليات حسابية قد نحتاج إليها) .
 ويقصد بالاستدلال الإحصائي عن تلك المعالم هو : (١) إنشاء فترات ثقة لتلك
 المعالم (٢) اختبارات المعنوية لكل معلمة من تلك المعالم . لكن هذا الاستدلال
 يقتضى معرفة توزيع المعاينة لتقديرات تلك المعالم ، أى توزيع المعاينة لكل من
 β_1 ، β_2 . ومنها يمكن استنتاج الخطأ المعياري لتلك التقديرات . عموماً وبدون
 الدخول فى تفاصيل رياضية ، ومع فرض تحقق شروط أو فروض نموذج
 الإنحدار المتعدد ، فإن تباين تقديرات طريقة المربعات الصغرى β_1 ، β_2 على
 الصورة التالية :

$$\text{ع}^2 (\beta_1) = \text{خ}^2_{\text{مر}} \times \frac{\text{مج} - (\text{س}_1 - \text{س}_2)^2}{\text{م}} \quad (21) \dots$$

$$\text{ع}^2 (\beta_2) = \text{خ}^2_{\text{مر}} \times \frac{\text{مج} - (\text{س}_1 - \text{س}_2)^2}{\text{م}} \quad (22) \dots$$

حيث :

$\text{خ}^2_{\text{مر}} = \text{مج} - (\text{ص} - \text{ن})^2 / 3$ أو الصورة المبسطة لها بالمعادلة رقم (١٣)
 $\text{م} = \text{مج} - (\text{س}_1 - \text{س}_2)^2 / 3 \times \text{مج} - (\text{س}_1 - \text{س}_2)^2 / 3 - [\text{مج} - (\text{س}_1 - \text{س}_2)^2 / 3]$
 أما $\text{ع} (\beta_1)$ ، $\text{ع} (\beta_2)$ فهي الأخطاء المعيارية لتلك التقديرات .

وعن طريق معرفة الأخطاء المعيارية $\text{ع} (\beta_1)$ ، $\text{ع} (\beta_2)$ يمكن إنشاء
 فترات الثقة للمعالم β_1 ، β_2 و إجراء اختبارات المعنوية لها فمثلاً ، فترة الثقة
 للمعلمة β_1 تكون على الصورة

$$\beta_1 = \beta_1 \pm \text{ت} (\alpha/2, \text{د} - 3) \times \text{ع} (\beta_1) \quad (23) \dots$$

بالمثل :

$$\beta_2 = \beta_2 \pm \text{ت} (\alpha/2, \text{د} - 3) \times \text{ع} (\beta_2) \quad (24) \dots$$

أما فيما يتعلق باختبار المعنوية للمعلمة β ، فتتم على النحو التالي :

١- الفرض العدمي : $\beta = 0$ صفر (لا يوجد تأثير للمتغير المستقل X_1 على التابع Y)

على التابع Y)

٢- الفرض البديل : $\beta \neq 0$ صفر يوجد تأثير للمتغير X_1 على التابع Y

$$3- \text{ت} = \frac{\beta - 0}{\text{ع}(\beta)} = \frac{\beta}{\text{ع}(\beta)} \quad , \text{حيث } \beta = 0 \text{ صفر}$$

٤- ت الجدولية : ت $\alpha/2$ ، ت $1-\alpha/2$

٥- المقارنة بين ت المحسوبة ، ت الجدولية .

٦- اتخاذ القرار بقبول أو رفض الفرض العدمي .

وتتبع نفس الخطوات عند اختبار $\beta = 0$ صفر .

مثال (٣)

مستخدماً بيانات مثال (١) ، المطلوب

(أ) فترة ثقة للمعلمة β عند $\alpha = 5\%$

(ب) اختبار معنوية المعلمة β عند $\alpha = 5\%$

الحل

(أ) فترة الثقة ٩٥% للمعلمة β :

$$\beta \pm \text{ت} \times \text{ع}(\beta) \quad \text{ت} = 1,96 \quad , \quad \text{ع}(\beta) = 0,00025$$

$$\text{حيث } \beta = 4,303 \quad , \quad \text{ت} = 1,96 \quad , \quad \text{ع}(\beta) = 0,00025$$

أما الخطأ المعياري $\text{ع}(\beta)$ وفقاً للمعادلة (٢٢) :-

$$\text{ع}(\beta) = \sqrt{\frac{\text{مجم س}^2 - \frac{(\text{مجم س})^2}{\text{ن}}}{\text{م}}}$$

حيث مجم س^2 الحصول عليها = ٠,٨٦٦ أما مجم س فهي :

$$\text{م} = \text{مجم س}^2 - \frac{(\text{مجم س})^2}{\text{ن}} = 0,866 - \frac{(0,866)^2}{10} = 0,866 - 0,075 = 0,791$$

$$- [\text{مجم س } ٢ - \text{مجم س } ١ \times \text{مجم س } ٢ / \text{ن}]^2$$

$$م = ٥٥ - [٥ / (١٥)] \times [٥ / (٢٥) - ١٢٩] - [٥ / (٢٥) - ٨١] \times [٥ / (٢٥) - ١٥] = ٤$$

$$- [٤ \times ١٠] - [٤ \times ١٠] = ٤ - ٣٦ - ٤٠ = ٦$$

$$\therefore \text{ع (ب) } = ٠,٨٦٦ \times \sqrt{\frac{١٠}{٤}} = ١,٣٦٩ = ١,٥٨١ \times ٠,٨٦٦$$

$$\therefore \beta = ١,٥ - ٤,٣٠٣ \times ١,٣٦٩ = ٥,٨٩ \pm ١,٥٠$$

$$= ٧,٣٩ - ٤,٣٩ +$$

(ب) إختبار معنوية β = صفر

١- الفرض العدمي β = صفر

٢- الفرض البديل $\beta \neq$ صفر

$$٣- ت = \frac{\beta - \beta}{\text{ع (ب) } (١)} = \frac{\beta}{\text{ع (ب) } (١)}$$

حيث $\beta = ٢,٥$ أما $\text{ع (ب) } (١)$ فهي

$$\text{ع (ب) } (١) = \text{مجم س } ٢ - \frac{(\text{مجم س } ١)^2}{\text{ن}} = ٤ - \frac{(١٠)^2}{٢٥} = ٠,٨٦٦$$

$$\therefore ت = \frac{٢,٥}{٠,٨٦٦} = ٢,٨٨$$

$$٤- ت الجدولية = ت (ن-٣, α / ٢) = ت (٢٥-٣, ٠,٠٢٥ / ٢) = ٤,٣٠٣$$

٥- المقارنة ت المحسوبة أقل من ت الجدولية

٦- القرار : قبول الفرض العدمي

$\therefore \beta =$ صفر ، أى لا يوجد تأثير معنوى للمتغير س١ على التابع ص

ملحوظة :

من الممكن أن تستخدم فترة الثقة للمعلة β ، أو β فى إختبار

$\beta =$ صفر أو $\beta \neq$ صفر وذلك بالبحث عن وجود أو عدم وجود الصفر داخل

حدى السقة . فإذا وقع الصفر داخل حدى الثقة ، لكان القرار قبول الفرض العدمى بأن $\beta_1 = 0$. صفر مثلاً أما إذا وقع الصفر خارج حدى الثقة ، لكان القرار هو رفض الفرض العدمى أى قبول الفرض البديل بأن $\beta_1 \neq 0$.

(٨) إختبار المعنوية الكلية للإنحدار: مدغل تحليل التباين

Analysis Of Variance

أحد أهداف تحليل نموذج الانحدار ، هو تحديد ما إذا كانت المتغيرات التفسيرية الداخلة فى النموذج ككل أى مجتمعة مع بعض ، ذات علاقة إرتباط حقيقية بالمتغير التابع ص أم لا . لنفرض أنه لا توجد علاقة حقيقية بين التابع ص و المتغيرات التفسيرية أو المستقلة س_١ ، س_٢ . هذا الفرض يعنى أن معاملات الانحدار فى مجتمع الدراسة $\beta_1, \beta_2, \dots, \beta_k = 0$. صفر . مقابل ذلك يظهر الفرض البديل الذى يعنى أن هناك أحد المتغيرات المستقلة - أن لم يكن كلها - على الأقل ذو علاقة حقيقية بالمتغير التابع ص أى $\beta_1 \neq 0, \beta_2 \neq 0, \dots, \beta_k \neq 0$.

بنفس الأسلوب الذى اتبع فى حالة استخدام تحليل التباين فى الانحدار البسيط ، فإننا نقسم الاختلافات الكلية فى التابع ص إلى جزئين (١) تغيرات أو اختلافات ناتجة عن وجود المتغيرات التفسيرية س_١ ، س_٢ (٢) تغيرات أو اختلافات ناتجة عن عوامل عشوائية أو ما تسمى أحياناً بالبقاوى . أى أن :

مجموع مربعات الاختلافات الكلية = مجموع المربعات المفسر بوجود

المتغيرات المستقلة + مجموع المربعات المفسر (البقاوى)

وعندما نقترن مجموع المربعات بدرجات الحرية ، نحصل على التباين ، أو متوسط مجموع المربعات . وعن طريق قسمة متوسط مجموع المربعات المفسر على متوسط مجموع المربعات غير المفسر (البقاوى) نحصل على متغير عشوائى (ف) ، وهو متغير يتبع توزيع ف بدرجات حرية (ك-١) ، (ن-ك) ، حيث ك : عدد المعالم المقدره . بمقارنة قيمة ف مع قيمة جدولية من توزيع ف عند درجتى حرية (ك-١) ، (ن-ك) ، α يمكن أى نصل إلى قرار بقبول أو

رفض الفرض $\beta_1 = \beta_2 = \text{صفر}$.

وبترتيب مجاميع المربعات السابقة داخل جدول تحليل التباين - بنفس الأسلوب الذي اتبع في حالة الانحدار البسيط - مقرونة بدرجات الحرية ، نصل إلى الأحصاء (ف) على الصورة التالية:

$$\text{ف} = \frac{\text{مجموع المربعات المفسر} \div (\text{ك} - 1)}{\text{مجموع المربعات غير المفسر} \div (\text{ن} - \text{ك})}$$

$\frac{\text{إِب، مَج من س، مَج من خ} \times \text{مَج من ا(ن)، ب، مَج من س، مَج من عَج من ز، ن(ك) - (ا)}{\text{مَج خ} = \text{مَج (من - ا)} + (\text{ن - ك})}$

حيث ك : عدد المعالم المقدرة : ب ، ب ، ب ، ن : حجم العينة
جدير بالذكر أن هناك صورة رياضية بديلة لصورة ف السابقة ، وهي تعتمد
على معلومية معامل التحديد المتعدد R^2 ، ومن السهل الوصول إليها وهي على
الصورة :

$$(٢٥) \dots \boxed{\frac{\text{ن-ك}}{١-ك} \times \frac{\text{ر}}{\text{ر}-١}} = \frac{\text{ر} \div (\text{ك}-١)}{(\text{ن}-\text{ك}) \div (\text{ر}-١)} = \text{ف}$$

مثال (۴) :

مستخدما بيانات مثال (١)

المطلوب اختبار $\beta_1 - \beta_2$ - صفر عند $\alpha = 5\%$ ، مستخدما في ذلك أسلوب تحليل التباين .

الحل :

يقتضى أسلوب تحليل التباين حساب المكونات التالية :

مجموع المربعات الكلي في المتغير التابع ص -

م.م.ك - مج ٢ ص ٢ - (مج ٢ ص) ٢/ن

$$\gamma_A = \lambda_0 - 1 \cdot \lambda = 0 \quad / \quad \gamma(\gamma_0) = 1 \cdot \lambda =$$

مجموع المربعات المفسر بوجود س_١، س_٢

$$= \text{ب} - [\text{مـجـ ص س} - \text{مـجـ ص} \times \text{مـجـ س} / \text{ن}]$$

$$+ \text{ب} - [\text{مـجـ ص س} - \text{مـجـ ص} \times \text{مـجـ س} / \text{ن}]$$

$$\text{م.م. بسبب الانحدار} = ٢,٥ - [٥/١٥ \times ٢٠ - ٧٦] ١,٥ - [٥/٢٥ \times ٢٠ - ١٠٩] ١,٥ - ١٣,٥ - ٤٠ = ٩ \times ١,٥ - ١٦ \times ٢,٥ =$$

$$٢٦,٥ = ١٣,٥ - ٤٠ = ٩ \times ١,٥ - ١٦ \times ٢,٥ =$$

جدول تحليل التباين

المصدر	م.م	د. ح	م.م	ف
م.م. بسبب انحدار	٢٦,٥	ك-١ = ٢	١٣,٢٥	
ص/س، س _٢				١٧,٦٦٧
م.م. د. (البواقي)	١,٥	ن-ك = ٢	٠,٧٥ = خ ^٢ ص	
م.م. ك	٢٨	ن-١ = ٤		

خطوات الاختبار :

١- الفرض العدمي : $\beta_1 = \beta_2 = ٠$ صفر (لا توجد علاقة بين ص وكل

من س_١، س_٢)

٢- الفرض البديل : $\beta_1 \neq ٠, \beta_2 \neq ٠$ صفر (توجد علاقة خطية من

المتغيرات الثلاث)

$$١٧,٦٦٧ = \text{ف} - ٣$$

$$١٩ = \text{ف الجدولية : ف} (٢,٢٠٥\%) = ١٩$$

٥- المقارنة : ف المحسوبة أقل من ف الجدولية .

٦- القرار : قبول الفرض العدمي .

∴ $\beta_1 = \beta_2 = ٠$ صفر أي لا توجد علاقة حقيقية بين ص وكل من س_١،

س_٢ . ومن ثم فإن هذه العلاقة لا تصلح في عملية التنبؤ بالمتغير التابع

ص .

ملحوظة :

من الممكن استخدام معامل التحديد R^2 ومعامل عدم التحديد $(1-R^2)$ في قياس قيمة F باستخدام العلاقة (٢٥) على النحو التالي .

$$F = \frac{R^2}{1-R^2} \times \frac{n-k}{k-1} , \text{ حيث } k=3, n=5$$

فإذا كانت R^2 معلومة من بيانات سابقة [$R^2 = 26.5 \div 28 = 0.9464$] ، فإن قيمة F تصبح :

$$F = \frac{0.9464}{1-0.9464} \times \frac{3-5}{5-3} = \frac{0.9464}{0.0536} \times \frac{-2}{2} = -17.667$$

وهي نفس القيمة التي ظهرت في جدول تحليل التباين .

(٩) اختبار مساهمة المتغيرات التفسيرية : مبدأ مجموع المربعات الإضافي

The Extra Sum of Squares Principle

بيننا من قبل كيف يمكن الاعتماد على أسلوب تحليل التباين في اختبار العلاقة الكلية لنموذج الانحدار ، أي اختبار مدى وجود علاقة حقيقية بين المتغيرات التفسيرية ككل (S_1, S_2) مع المتغير التابع Y أي اختبار $\beta_1 = \beta_2 = 0$. أيضا بيننا من قبل كيف يتم اختبار مدى وجود علاقة حقيقية بين كل متغير تفسيري على حدة S_1, S_2 والمتغير التابع Y ، أي اختبار $\beta_1 = 0$ ، $\beta_2 = 0$ واختبار $\beta_1 = 0$ ، $\beta_2 = 0$. في اختبار العلاقة الإجمالية ، استخدم توزيع F ، وفي اختبار كل متغير على حدة ، استخدم توزيع t . غير أن هناك وسيلة أخرى لاختبار معنوية β_1 أو اختبار معنوية β_2 وذلك بالبحث في مدى أهمية وجود المتغير S_1 أو المتغير S_2 في نموذج الانحدار ، بمعنى ، هل وجود المتغير S_1 له أهمية عند دراسة العلاقة بين S_1, S_2 أم يمكن إهماله وحذفه من نموذج الانحدار ، أيضا ، هل من المفيد والضروري أن يظهر المتغير S_2 في نموذج الانحدار أم يمكن إهماله وحذفه ؟ هذه الأسئلة سبق أن أجيب عنها عند

اختبار معنوية معاملات خط الانحدار باستخدام اختبار ت . لكن هناك وسيلة أخرى يمكن الاستعانة بها لاختبار معنوية المتغيرات التفسيرية كل على حدة ، تعتمد على أسلوب تحليل التباين والفكرة الأساسية هنا يمكن إيضاحها كالآتي :

(١) إذا كنا نبحث في مدى أهمية وجود المتغير المستقل x_1 مثلاً في نموذج الانحدار ، فإننا نقارن بين مجموع المربعات بسبب وجود x_1 معاً ومجموع المربعات بسبب وجود x_2 ، والفرق بين المجموعين لا بد وأن يرجع إلى مجموع المربعات الإضافي Extra sum of squares للمتغير x_1 أي الزيادة في مجموع المربعات بسبب وجود x_1 . أي أننا نقارن بين نموذجين للانحدار ، الأول بدلالة x_1 ، والثاني بدلالة x_2 فقط أي :

$$\text{النموذج الأول : ص} = \text{ب} + \text{ب} \cdot \text{س}_1 + \text{س}_2 \cdot \text{س}_2$$

$$\text{النموذج الثاني : ص} = \text{ب} + \text{ب} \cdot \text{س}_2 + \text{س}_2$$

أما مجموع المربعات في ظل النموذج الأول والنموذج الثاني فهي على الصورة

$$\text{م.م. بسبب انحدار ص/س}_1 = \text{ب} \cdot \text{ب} - (\text{م.م. ص} - \text{م.م. ص} \cdot \text{م.م. س}_1 / \text{ن})$$

$$+ \text{ب} \cdot \text{ب} - (\text{م.م. ص} - \text{م.م. ص} \cdot \text{م.م. س}_2 / \text{ن})$$

$$\text{م.م. بسبب انحدار ص/س}_2 = \frac{(\text{م.م. ص} - \text{م.م. ص} \cdot \text{م.م. س}_1 / \text{ن}) - (\text{م.م. ص} - \text{م.م. ص} \cdot \text{م.م. س}_2 / \text{ن})}{\text{م.م. ص} - \text{م.م. ص} \cdot \text{م.م. س}_2 / \text{ن}}$$

(٢٦)...

وإذا كان م.م. بسبب انحدار ص / س₁ له (٢) درجة حرية ، فإن م.م.

بسبب انحدار ص / س₂ ، يكون له درجة حرية واحدة .

وبترتيب تلك المجاميع مع مجموع المربعات الكلي ومجموع مربعات البواقي أو

الخطأ العشوائي في جدول تحليل التباين ، يمكن اختبار مدى أهمية وجود x_1

في نموذج الانحدار أي اختبار معنوية $\beta_1 = 0$ صفر وذلك بإيجاد قيمة الإحصاء

ف حيث :

$$\text{ف} = \frac{\text{مجموع المربعات الإضافي أي مساهمة س}_1}{\text{تباين البواقي} - \text{س}_2} \quad (٢٧) \dots$$

وبمقارنة قيمة F السابقة مع قيمة جدولية من توزيع F بدرجات حرية ١ ،
(ن - ك) يمكن اتخاذ قرار بقبول أو رفض الفرض العدمي بأن $\beta_1 = 0$ - صفر .
بالطبع وكما هو متوقع لا بد وأن تتفق نتيجة القرار هنا مع القرار الذي يمكن أن
نصل إليه لو استخدمنا اختبار (ت) في اختبار معنوية $\beta_1 = 0$ - صفر الذي أشرنا
إليه سابقا .

(٧) إذا كنا نبحث في مدى أهمية وجود المتغير المستقل X_2 في نموذج
الانحدار ، فإننا نقارن بين مجموع المربعات في ظل وجود X_1, X_2 مع مجموع
المربعات في ظل وجود X_1 فقط ، والفرق بين المجموعين يرجع إلى مجموع
المربعات الإضافي للمتغير X_2 أي مساهمة X_2 في مجموع المربعات بسبب
انحدار Y على X_1, X_2 . بمعنى آخر فإننا نقارن بين نموذجين ، الأول بدلالة X_1 ،
 X_2 ، والثاني بدلالة X_1 فقط أي :

النموذج الأول : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

النموذج الثاني : $Y = \beta_0 + \beta_1 X_1 + \epsilon$

أما مجموع المربعات بسبب انحدار Y على X_1, X_2 فقد سبق ذكره ، في حين أن
مجموع المربعات بسبب انحدار Y على X_1 فهو :

$$M.م. انحدار Y على X_1 = \frac{(م.م. Y - م.م. X_1 \times م.م. Y / N)}{م.م. X_1 - (م.م. X_1)^2 / N}$$

وبترتيب تلك المجاميع في جدول تحليل التباين مع م.م. ك ، خ ، م ، يمكن
إيجاد مجموع المربعات الإضافي للمتغير X_2 والذي يتم اختباره على النحو
التالي :

$$F = \frac{\text{مجموع المربعات الإضافي أي مساهمة } X_2}{\text{تباين البواقي}}$$

وبالمقارنة مع F الجدولية عند درجتَي حرية (١ ، ن - ك) ، α يمكن أن
نصل إلى قرار بقبول أو رفض الفرض العدمي $\beta_2 = 0$ - صفر

مثال (٥)

استخدم بيانات مثال (١) والنتائج التي توصلنا إليها في الأمثلة التالية له والمرتبطة به في إجراء :

(١) اختبار $\beta_1 = 0$ - صفر عند $\alpha = 5\%$ مستخدماً أسلوب تحليل التباين .

(٢) اختبار $\beta_2 = 0$ - صفر عند $\alpha = 5\%$ مستخدماً أسلوب تحليل التباين .

الحل :

اختبار $\beta_1 = 0$ - صفر ، اختبار $\beta_2 = 0$ - صفر ، سبق أن تعرضنا لها عند الحديث عن اختبار معنوية معالم خط الانحدار باستخدام توزيع ت ومثال (٣) كان تطبيقاً لذلك . الآن نعيد تلك الاختبارات لكن عن طريق أسلوب تحليل التباين على النحو التالي :

(١) اختبار $\beta_1 = 0$ - صفر

لتكوين محتوى جدول تحليل التباين ، نجري الآتي :

م.م. ك - مج ص - ٢ - (مج ص) / ٢ / ن = ٢٨ (كما سبق)

م.م. بسبب انحدار ص/س ، ١ ، ص = ٢٦,٥ (كما سبق)

م.م. بسبب انحدار ص/س ، ٢ = $\frac{\text{مج ص} - \text{مج ص} \times \text{مج ص} / \text{ن}}{\text{مج ص} - \text{مج ص} / \text{ن}}$

$$٢٠,٢٥ = \frac{٨١}{٤} = \frac{١٠٩ - ٥/٢٥ \times ٢٠}{٥ - ٢٥/٢٥} = \frac{١٢٩ - ٢٠}{٥ - ١} = ٢٠,٢٥$$

جدول تحليل التباين لاختبار $\beta_1 = 0$ - صفر

المصدر	م.م	د. ح	م.م	ف = نسبة التباين
م.م. انحدار ص/س	٢٠,٢٥	١	٦,٢٥	$\frac{٦,٢٥}{٠,٧٥} = ٨,٣٣$
الزيادة بسبب ص ، أو مساهمة ص	٦,٢٥	١		
م.م. انحدار ص/س ، ٢	٢٦,٥	٢	٠,٧٥	
م.م. د	١,٥	٢		
م.م. ك	٢٨	٤		

خطوات الاختبار :

- ١- الفرض العدمي $\beta_1 = 0$ صفر (لا يوجد تأثير للمتغير x_1 على التابع y أي يمكن استبعاده من نموذج الانحدار)
 - ٢- الفرض البديل $\beta_1 \neq 0$ صفر
 - ٣- ف المحسوبة = ٨,٣٣
 - ٤- ف الجدولية ، ف (١,٢,٥) = ١٨,٥١
 - ٥- المقارنة ، ف المحسوبة أقل من ف الجدولية
 - ٦- القرار : قبول الفرض العدمي بأن $\beta_1 = 0$ صفر
- أي أن وجود x_1 في معادلة خط الانحدار غير مفيد في عملية التنبؤ بالمتغير y وهذا القرار سبق أن توصلنا إليه في مثال (٣) .

(٢) اختبار $\beta_2 = 0$ صفر

م.م.ك = ٢٨ ، م.م. بسبب انحدار y / x_1 و x_2 = ٢٦,٥

$$F = \frac{(م.م. - م.م. ك) / (م.م. - م.م. ك) / (م.م. - م.م. ك)}{(م.م. - م.م. ك) / (م.م. - م.م. ك) / (م.م. - م.م. ك)}$$

$$F = \frac{(26.5 - 28) / (2 - 1)}{(26.5 - 28) / (2 - 1)} = \frac{-1.5}{-1.5} = 1$$

جدول تحليل التباين لاختبار $\beta_2 = 0$ صفر

المصدر	م.م	د.ح	م.م.م	نسبة التباين ف
م.م. انحدار y / x_1	٢٥,٦	١	٠,٩	$1,2 = \frac{0,9}{0,75}$
مساهمة x_1	٠,٩	١		
م.م. انحدار y / x_2	٢٦,٥	٢	٠,٧٥	
م.م.د (البواقي)	١,٥	٢		
م.م.ك	٢٨	٤		

خطوات الاختبار :

- ١- الفرض العدمي $\beta_2 = 0$ - صفر
 - ٢- الفرض البديل $\beta_2 \neq 0$ - صفر
 - ٣- ف المحسوبة = ١,٢
 - ٤- ف (١٠,٢٥%) = ١٨,٥١
 - ٥- المقارنة : ف المحسوبة أقل من ف الجدولية
 - ٦- القرار : قبول الفرض العدمي
- ∴ $\beta_2 = 0$ - صفر ، أي أن وجود س_٢ لا يفيد معنوياً في تفسير التغير في ص وبالتالي لا يفيد معنوياً في التنبؤ بقيم التابع ص .

مثال (٦) (شامل) :

الجدول التالي يعطي كمية الإنتاج من القمح بالطن للفدان (ص) وكمية السماد بالكيلو جرام (س_١) وكمية المبيدات الحشرية بالرطل (س_٢) في إحدى المزارع خلال عشر سنوات متتالية من ١٩٩١ إلى ٢٠٠٠

السنة	١٩٩١	٩٢	٩٣	٩٤	٩٥	٩٦	٩٧	٩٨	٩٩	٢٠٠٠
ص	٤٠	٤٤	٤٦	٤٨	٥٢	٥٨	٦٠	٦٨	٧٤	٨٠
س _١	٦	١٠	١٢	١٤	١٦	١٨	٢٢	٢٤	٢٦	٣٢
س _٢	٤	٤	٥	٧	٩	١٢	١٤	٢٠	٢١	٢٤

المطلوب :

- ١- إيجاد معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى .
- ٢- الخطأ المعياري لتقدير معادلة الانحدار .
- ٣- معامل التحديد ومعامل التحديد المعدل ومعامل الارتباط المتعدد .
- ٤- معاملات الارتباط الجزئية ثم حدد أي المتغيرات المستقلة يساهم أكثر في تفسير التغير في التابع ص .

٥- قدر ١٣ بفترة ثقة ٩٥%

٦- اختبار β_1 - صفر ، اختبار β_2 - صفر عند $\alpha = 5\%$ مستخدما توزيع ت

٧- اختبار β_1 = صفر ، اختبار β_2 = صفر عند $\alpha = 5\%$ مستخدماً توزيع ف

(تحليل التباين)

٨- اختبار المعنوية الاجمالية للانحدار المتعدد أي اختبار $\beta_1 = \beta_2 = \dots = \beta_k = 0$ - صفر

مستخدماً أسلوب تحليل التباين ($\alpha = 5\%$)

الحل :

م	ص	ص	ص	ص ^٢	ص ^٢	ص ^٢	ص ^٢	ص ^٢	ص ^٢
١	٤٠	٦	٤	١٦٠٠	٣٦	١٦	٢٤٠	١٦٠	٢٤
٢	٤٤	١٠	٤	١٩٣٦	١٠٠	١٦	٤٤٠	١٧٦	٤٠
٣	٤٦	١٢	٥	٢١١٦	١٤٤	٢٥	٥٥٢	٢٣٠	٦٠
٤	٤٨	١٤	٧	٢٣٠٤	١٩٦	٤٩	٦٧٢	٣٣٦	٩٨
٥	٥٢	١٦	٩	٢٧٠٤	٢٥٦	٨١	٨٣٢	٤٦٨	١٤٤
٦	٥٨	١٨	١٢	٣٣٦٤	٣٢٤	١٤٤	١٠٤٤	٦٩٦	٢١٦
٧	٦٠	٢٢	١٤	٣٦٠٠	٤٨٤	١٩٦	١٣٢٠	٨٤٠	٣٠٨
٨	٦٨	٢٤	٢٠	٤٦٢٤	٥٧٦	٤٠٠	١٦٣٢	١٣٦٠	٤٨٠
٩	٧٤	٢٦	٢١	٥٤٧٦	٦٧٦	٤٤١	١٩٢٤	١٥٥٤	٥٤٦
١٠	٨٠	٣٢	٢٤	٦٤٠٠	١٠٢٤	٥٧٦	٢٥٦٠	١٩٢٠	٧٦٨
مج	٥٧٠	١٨٠	١٢٠	٣٤١٢٤	٣٨١٦	١٩٤٤	١١٢١٦	٧٧٤٠	٢٦٨٤

(١) معادلة خط الانحدار المتعدد : $ص = ب + ب١ س١ + ب٢ س٢$ عن طريق

المعادلات الطبيعية الثلاثة التالية ، يمكن إيجاد التقديرات ب ١ ، ب ٢ كمايلي :

مجـ ص = ن ب + ب ۱ مجـ س ۱ + ب ۲ مجـ س ۲

$$\text{مجموع ص ۱} = \text{ب مجموع ص ۱} + \text{ب ۱ مجموع ص ۱} + \text{ب ۲ مجموع ص ۱} + \text{ب ۳ مجموع ص ۱}$$

مجس ص ۲ ب مجس ۲ + ب ۱ مجس ۱ س ۱ س ۲ + ب ۲ مجس ۲ ۲

وبالتعويض عن المجاميع السابقة بالقيم الموجودة بالجدول ، حيث $n = 10$
 نحصل على ثلاث معادلات في ثلاث مجاهيل ب ، ب_١ ، ب_٢
 $57. = 10 \text{ ب} + 180 \text{ ب}_1 + 120 \text{ ب}_2$
 $11216 = 180 \text{ ب} + 3816 \text{ ب}_1 + 2684 \text{ ب}_2$
 $7740 = 120 \text{ ب} + 2684 \text{ ب}_1 + 1944 \text{ ب}_2$

وبحل تلك المعادلات بأي طريقة جبرية كالحذف أو المحددات أو المصفوفات
 نحصل على القيم التالية : $\text{ب} = 31,98$ ، $\text{ب}_1 = 0,65$ ، $\text{ب}_2 = 1,11$
 ∴ معادلة خط الانحدار المتعدد تصبح على الصورة .

$$\text{ص} = 31,98 + 0,65 \text{ س}_1 + 1,11 \text{ س}_2$$

بفحص معادلة الانحدار التي حصلنا عليها نجد أن تغير س_١ (كمية السماد) بوحدة واحدة يؤدي إلى تغير ص (كمية الإنتاج من القمح) بمقدار ٠,٦٥ وحدة ، وذلك بغض النظر عن قيمة س_٢ (كمية المبيدات الحشرية) ، وينطبق نفس المفهوم بالنسبة لمعامل الانحدار الجزئي ب_٢ = ١,١١ ، بمعنى أن تغير كمية المبيدات الحشرية بوحدة واحدة يؤدي إلى تغير كمية الإنتاج من القمح (ص) بمقدار ١,١١ وحدة وذلك بغض النظر عن قيمة س_١ .
 إن إحدى المشاكل الهامة التي تواجه استخدام الانحدار المتعدد ، هي إمكانية وجود ارتباط بين المتغيرات المستقلة ، وتسمى هذه المشكلة بالازدواج الخطي بين المتغيرات المستقلة Multicollinearity ، فإذا كان هذا الارتباط قويا ، أدى ذلك إلى تضالٍ مصداقية معاملات الانحدار الجزئية ، وسوف نتناول هذه المشكلة في مرحلة لاحقة عند الحديث عن مشاكل تطبيق الانحدار المتعدد .

(٢) الخطأ المعياري لتقدير معادلة الانحدار

$$\text{خطأ معياري} = \sqrt{\frac{\text{مجم ص} - \text{مجم ص}^2 - \text{ب} \cdot \text{مجم ص} - \text{ب}_1 \cdot \text{مجم ص} - \text{ب}_2 \cdot \text{مجم ص}}{(n-3)}}$$

$$\sqrt{\frac{(3-10) \div (7740 \times 1,11 - 11216 \times 0,65 - 570 \times 31,98 - 34124)}{1,394 - 1,943 - 7 \div 13,6}} = \sqrt{\frac{7 \div [34110,4 - 34124]}{1,394 - 1,943 - 7 \div 13,6}} =$$

(3) معامل التحديد ومعامل التحديد المعدل : [معادلة (8) ، معادلة (11)]
معامل التحديد $R^2 =$

$$\frac{\text{ب} \cdot [\text{مجم من س} - 1 - \text{مجم من س} \times \text{مجم من س} / \text{ن}] + \text{ب} \cdot [\text{مجم من س} - 2 - \text{مجم من س} \times \text{مجم من س} / \text{ن}]}{\text{مجم من س} - 2 - (\text{مجم من س} / \text{ن})} = R^2$$

$$R^2 = \frac{[10 / 120 \times 570 - 7740] 1,11 + [10 / 180 \times 570 - 11216] 0,65}{10 / (570) - 34124} =$$

$$1634 \div 1620,4 = 1634 \div (900 \times 1,11 + 906 \times 0,65) = 0,9917 =$$

أي أن 99,17% من التغيرات الكلية في إنتاج القمح تفسر بسبب وجود كل من السماد (س1) والمبيدات (س2) أما الباقي 0,83% فنرجع إلى عوامل عشوائية غير مفسرة .

معامل الارتباط المتعددة هو الجذر التربيعي لمعامل التحديد .

$$R = \sqrt{0,9917} = 0,9958$$

معامل التحديد المعدل R^2 والذي يأخذ في الاعتبار النقص في درجات الحرية عند إضافة متغيرات مستقلة لنموذج الانحدار يأخذ شكل المعادلة رقم (11) السابق نذكرها .

$$R^2 = 1 - \frac{1 - \text{ن}}{\text{ن} - \text{ك}} \times (2 - 1)$$

$$= 1 - \frac{1 - 10}{3 - 10} \times (0,9917 - 1) = \frac{9}{7} \times 0,0083 - 1 =$$

$$= 1 - 0,107 = 0,893 = 89,3\%$$

بالطبع قيمة R^2 هي دائما أقل من R^2

(٤) معاملات الارتباط الجزئية

يمكن حساب معاملات الارتباط الجزئية بدلالة معاملات الارتباط الخطية البسيطة على النحو التالي :

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2} \sqrt{1 - r_{23}^2}}$$

$$r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{1 - r_{12}^2} \sqrt{1 - r_{23}^2}}$$

$$r_{23.1} = \frac{r_{23} - r_{12}r_{13}}{\sqrt{1 - r_{12}^2} \sqrt{1 - r_{13}^2}}$$

حيث r_{12} ، r_{13} ، r_{23} هي معاملات الارتباط الخطية البسيطة حيث الأرقام ١، ٢، ٣ تعني x_1 ، x_2 ، x_3 تعني y

$$r_{12.3} = \frac{\text{مجد ص ١} - \text{مجد ص ٢} \times \text{مجد ص ٣} / \text{ن}}{\sqrt{[\text{مجد ص ١} - (\text{مجد ص ٢})^2 / \text{ن}] [\text{مجد ص ٣} - (\text{مجد ص ٣})^2 / \text{ن}]}}$$

$$= \frac{10 / 180 \times 570 - 11216}{\sqrt{[10 / 2(180) - 3816] [10 / 2(570) - 32124]}}$$

$$= \frac{956}{576 \times 1634} = \frac{956}{970.1} = 0.9854$$

رمر ٢ : بوضع x_2 مكان x_1 في العلاقة السابقة والتعويض نجد أن

$$r_{13.2} = \frac{10 / 120 \times 570 - 7740}{\sqrt{[10 / 2(120) - 1944] \times 1634}}$$

$$= \frac{900}{504 \times 1634} = \frac{900}{907.49} = 0.9917$$

$$r_{23.1} = \frac{\text{مجد ص ١} - \text{مجد ص ٢} \times \text{مجد ص ٣} / \text{ن}}{\sqrt{[\text{مجد ص ١} - (\text{مجد ص ٢})^2 / \text{ن}] [\text{مجد ص ٣} - (\text{مجد ص ٣})^2 / \text{ن}]}}$$

وبالتعويض مع الاستعانة بالنتائج فى العلاقات السابقة :

$$r_{11} = \frac{0.9725}{0.9725} = \frac{10/120 \times 180 - 2684}{0.4 \times 576} = 0.9725$$

وبالتعويض عن r_{11} ، r_{12} ، r_{22} فى معاملات الارتباط الجزئية نصل إلى :

$$r_{12.1} = \frac{0.9725 \times 0.9917 - 0.9854}{0.9458 - 1} = 0.9725$$

$$r_{22.1} = \frac{0.9725 \times 0.9854 - 0.9917}{0.9458 - 1} = 0.9725$$

وحيث أن $r_{12.1}$ أكبر من r_{11} ، فهذا يعنى أن المتغير x_2 (كمية المبيدات الحشرية) أكثر أهمية من x_1 (كمية السماد) فى تفسير التغيرات فى إنتاج القمح (ص) .

(٥) تقدير β بفترة ثقة ٩٥ % :

$$\beta = b_1 \pm t_{(n-2)/2} \times \text{ع(ب)}$$

$$\text{حيث : } b_1 = 0.65 , \quad t_{(n-2)/2} = 2.365$$

أما الخطأ المعياري للتقدير (ب) أى ع (ب) فهو على شكل المعادلة

(٢١) أى :

$$\begin{aligned} \text{ع(ب)} &= \frac{\text{مجم } x_2^2 - (\text{مجم } x_2)^2 / n}{\text{م}} \times \frac{\text{مجم } x_2^2 - (\text{مجم } x_2)^2 / n}{\text{م}} \\ \text{م} &= (\text{مجم } x_1^2 - (\text{مجم } x_1)^2 / n) (\text{مجم } x_2^2 - (\text{مجم } x_2)^2 / n) - \\ &\quad - (\text{مجم } x_1 x_2 - (\text{مجم } x_1)(\text{مجم } x_2) / n) \\ &= [10/120 \times 180 - 2684] - \\ &\quad - [10/120 \times 180 - 2684] \end{aligned}$$

$$10728 = 274576 - 2903.4 = 2[524] - [50.4 \times 576] =$$

أما χ^2 من فقد سبق حسابها وكانت : $\chi^2 = 1.943$

$$\therefore \text{ع (ب)} = \frac{(50.4)}{10728} \times 1.943 = 0.006$$

∴ الخطأ المعياري ع(ب) = 0.24

$$0.8, 1.22 = 0.07 + 0.65 = 0.24 \times 2.365 + 0.65 = 0.6$$

وبالمثل يمكن إيجاد فترة الثقة للمعلمة β بعد حساب الخطأ المعياري

للتقدير ب، وهو على الصورة :

$$\text{ع (ب)} = \chi^2 \times \frac{(\text{مجم س } 1 - (\text{مجم س } 1) / \text{ن})}{\text{م}}$$

$$0.07 = \frac{576}{2(524) - 50.4 \times 576} \times 1.943 =$$

$$\therefore \text{ع (ب)} = 0.07 \Rightarrow 0.26$$

(٦) اختبار β = صفر واختبار β = صفر باستخدام توزيع ت :

أولاً : اختبار β = صفر

خطوات الاختبار :

١- الفرض العدمي β = صفر

٢- الفرض البديل β ≠ صفر

$$3- \text{ت} = \frac{\text{ب} - 1\beta}{\text{ع (ب)}} = \frac{0.65}{0.24} = 2.71$$

$$4- \text{ت الجدولية} : \text{ت} = \frac{\text{ت}}{\sqrt{50.4 \times 2}} = 2.365$$

٥- المقارنة : ت المحسوبة أكبر من ت الجدولية

٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .

∴ β_1, β_2 صفر (أي أن هناك علاقة حقيقية بين المتغير المستقل X_1 والمتابع Y ومن المهم وجود هذا المتغير في نموذج خط الانحدار المتعدد).

ثانياً: اختبار $\beta_2 = \text{صفر}$

خطوات الاختبار:

١- الفرض العدمي $\beta_2 = \text{صفر}$

٢- الفرض البديل $\beta_2 \neq \text{صفر}$

$$٣- \text{ت المحسوبة} = \frac{\text{ب} - \beta_2 \text{ صفر}}{\text{ع}(\text{ب})} = \frac{١,١١ - ٠,٢٦}{\text{ع}(\text{ب})} = ٤,٢٧$$

$$٤- \text{ت الجدولية} : \text{ت} = ٠,٠٢٥, ٧ = ٢,٣٦٥$$

٥- المقارنة: ت المحسوبة أكبر من ت الجدولية

٦- القرار: رفض الفرض العدمي وقبول الفرض البديل ∴ $\beta_2 \neq \text{صفر}$

(٧) اختبار $\beta_1 = \text{صفر}$ ، اختبار $\beta_2 = \text{صفر}$ باستخدام تحليل التباين:

أولاً: اختبار $\beta_1 = \text{صفر}$ أو اختبار مدى أهمية وجود X_1 في معادلة خط

الانحدار

$$\text{م.م.ك} = \text{مج ص}^2 - (\text{مج ص})^2 / \text{ن} = ٣٤١٢٤ - ١٠ / (٥٧٠) = ١٦٣٤$$

$$\text{م.م. بسبب انحدار ص/س} = \text{ب}(\text{مج ص س} - \text{مج ص} \times \text{مج س} / \text{ن})$$

$$\text{ب}(\text{مج ص س} - \text{مج ص} \times \text{مج س} / \text{ن})$$

$$= ٠,٦٥ (١١٢١٦ - ١٠ / ١٨٠ \times ٥٧٠)$$

$$+ ١,١١ (٧٧٤٠ - ١٠ / ١٢٠ \times ٥٧٠)$$

$$= ٩٥٦ \times ٠,٦٥ + ١,١١ \times ٩٠٠$$

$$= ٦٢١,٤ + ٩٩٩ = ١٦٢٠,٤$$

$$\text{م.م. بسبب انحدار ص/س} = \frac{(\text{مج ص س} - \text{مج ص} \times \text{مج س} / \text{ن})}{\text{مج ص}^2 - (\text{مج ص})^2 / \text{ن}}$$

$$16.7,14 = \frac{81.000}{5.4} = \frac{1(9.0)}{10/1(120) - 1944} =$$

جدول تحليل التباين

المصدر	م.م	د. ح	م.م	نسبة التباين
م.م. انحدار ص/س ^٢	16.7,14	1	1	$\frac{13,26}{1,943} = 6,82$
مساهمة س ^٢	13,26	1	1	
م.م. انحدار ص/س ^٢ س ^٢	1620,40	2	2	
م.م. د. (البواقي)	13,6	7	7	$1,943 = \chi^2$ ص
م.م. ك	1634	9	9	

خطوات الاختبار :

- ١- الفرض العدمي $\beta = 0$ صفر .
 - ٢- الفرض البديل $\beta \neq 0$ صفر .
 - ٣- ف المحسوبة $= 6,82$.
 - ٤- ف الجدولية : ف (١ ، ٧ ، ٥%) = ٥,٥٩ .
 - ٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .
 - ٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .
- ∴ $\beta \neq 0$ صفر أى أنه من المهم بقاء س^٢ فى معادلة خط الانحدار وهو نفس القرار الذى توصلنا إليه باختبار ت .

ثانياً: اختبار β - صفر أو اختبار مدى أهمية وجود س^٢ فى معادلة خط الانحدار

$$م.م. ك = 1634$$

$$م.م. بسبب انحدار ص/س^٢ س^٢ = 1620,40$$

$$م.م. نسب انحدار ص/س^٢ = \frac{(م.م. ص س^٢ - م.م. ص \times م.م. س^٢ / ن)}{م.م. س^٢ - م.م. س^٢ / ن}$$

$$1086,7 = \frac{10/180 \times 570 - 11216}{10/180 - 3816} = \frac{1086,7}{576} = \frac{10/180 \times 570 - 11216}{10/180 - 3816}$$

جدول تحليل التباين

المصدر	م.م	د. ح	م.م	نسبة التباين
م.م. انحدار ص/س ^١	١٥٨٦,٧	١	١	
مساهمة م ^٢	٣٣,٧	١	١	١٧,٣٤ =
م.م. انحدار ص/س ^١ م ^٢	١٦٢٠,٤	٢	٢	
م.م. د. (البواقي)	١٣,٦	٧	٧	١,٩٤٣ = م ^٢ م
م.م. ك	١٦٣٤	٩	٩	

خطوات الاختبار :

- ١- الفرض العدمي $\beta_1 = \beta_2 = \text{صفر}$.
 - ٢- الفرض البديل $\beta_1 \neq \beta_2$.
 - ٣- ف المحسوبة = ١٧,٣٤ .
 - ٤- ف الجدولية : ف (١ ، ٧ ، ٥٠) = ٥,٥٩ .
 - ٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .
 - ٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .
- ∴ $\beta_1 \neq \beta_2$ وهو نفس القرار الذي توصلنا إليه باستخدام توزيع ت .
- (٨) اختبار المعنوية الكلية للانحدار : اختبار $\beta_1 = \beta_2 = \text{صفر}$
- م.م. ك = مج ص^٢ - (مج ص) / ن = ١٦٣٤ =
- م.م. بسبب انحدار ص/س^١ م^٢ = ١٦٢٠,٤ =

جدول تحليل التباين

المصدر	م.م	د. ح	م.م	نسبة التباين
م.م. انحدار من ١ من ٢	١٦٢٠,٤	٢	٨١٠,٢	
م.م. د. (البواقي)	١٣,٦	٧	١,٩٤٣ = χ^2 من ٧	٤١٦,٩٨
م.م. ك	١٦٣٤	٩		

خطوات الاختبار :

- ١- الفرض العدمي $\beta_1 = \beta_2 = \text{صفر}$.
 - ٢- الفرض البديل $\beta_1 \neq \beta_2 \neq \text{صفر}$.
 - ٣- ف المحسوبة = ٤١٦,٩٨ .
 - ٤- ف الجدولية : ف (٢ ، ٧ ، ٥%) = ٤,٧٤ .
 - ٥- المقارنة : ف المحسوبة أكبر من ف الجدولية .
 - ٦- القرار : رفض الفرض العدمي وقبول الفرض البديل .
- $\therefore \beta_1 \neq \beta_2 \neq \text{صفر}$ أى أن هناك تأثير معنوى حقيقى لكل من س١ ، س٢ على المتغير التابع ص .

ملحوظة :

من الممكن حساب قيمة ف بطريقة أخرى عن طريق علاقة ف مع معامل التحديد ، حيث :

$$ف = \frac{r^2}{1 - r^2} \times \frac{ن - ك}{ك - ١} \quad \text{حيث ك عدد المعالم المقدرة .}$$

وحيث أن r^2 سبق الحصول عليها (أو يمكن الحصول عليها الآن من جدول تحليل التباين) حيث:

$$r^2 = \frac{\text{م.م بسبب الانحدار}}{\text{م.م ك}} = \frac{١٦٢٠,٤}{١٦٣٤} = ٠,٩٩١٧$$

$$\therefore \text{ف} = \frac{7}{2} \times \frac{0.9917}{0.0083} = \frac{3-1.0}{1-3} \times \frac{0.9917}{0.9917-1}$$

وهي تقريباً مساوية للقيمة التي ظهرت في جدول تحليل التباين .

(١٠) بعض مشاكل استخدام تحليل الانحدار :

عند تقدير معالم نموذج خط الانحدار -سواء البسيط او المتعدد- باستخدام طريقة المربعات الصغرى ، كانت هناك مجموعة من الافتراضات يشترط تحققها كي نستخدم هذه الطريقة . هذه الفروض سبق أن تعرضنا لها ويمكن تلخيصها على النحو التالي :

- ١- القيمة المتوقعة للحدود العشوائية (خ) = صفر .
- ٢- تباين الحدود العشوائية مقدار ثابت ويساوى ٦ .
- ٣- الحدود العشوائية (خ) تتبع توزيع طبيعي .
- ٤- الحدود العشوائية غير مرتبطة ببعضها البعض ، بمعنى أن الخطأ العشوائي في فترة ما غير مرتبط بالخطأ أو بالحد العشوائي في فترة أخرى .
- ٥- الحدود العشوائية (خ) والمتغيرات المستقلة أو المفسرة غير مرتبطة ببعضها البعض .
- ٦- هناك فرض إضافي خاص بالانحدار المتعدد ، وهو عدم وجود ارتباط بين المتغيرات المستقلة أو ما يسمى بالازدواج الخطي ، لأنه لو كان هناك ارتباط تام بين المتغيرات المستقلة لاستحال تقدير معالم خط الانحدار المتعدد بطريقة المربعات الصغرى .

لاشك أن هذه الفروض قد تتوفر من الناحية العملية أو قد لا تتوفر كلها ، فإذا ما توفرت تلك الفروض صلت طريقة المربعات الصغرى للاستخدام وأعطت تقديرات ذات دقة إحصائية عالية ، أما إذا لم تتوفر هذه الفروض فإن

طريقة المربعات الصغرى لا تصلح للاستخدام ، ويجب البحث عن طريقة تقدير أخرى . ولكي نتأكد من تحقق تلك الفروض يجب علينا إجراء بعض الاختبارات الإحصائية . من الفروض التي إذا سقطت أو لم تتحقق ترتب عليها مشاكل في التقديرات التي نتحصل عليها ما يلي :

- ١- مشكلة عدم ثبات تباين الحد العشوائي Heteroscedasticity
- ٢- مشكلة الارتواج الخطي Multicollinearity
- ٣- مشكلة الارتباط الذاتي للبوالي Autocorrelation

وتعد هذه المشاكل من أكثر المشاكل ظهوراً في التطبيقات العملية ، لذا يجب على الباحث التحقق من خلو البيانات التي يعالجها عن تلك المشاكل قبل استخدام طريقة المربعات الصغرى في عملية تقدير معالم خط الانحدار البسيط أو المتعدد . ونظراً لما تتطلبه هذه المشاكل من معالجة رياضية متقدمة ، بجانب الإلمام بمستوى متقدم من أساليب التحليل الإحصائية ، فسوف نكتفي بعرض تلك المشاكل في نطاق مختصر ، مكتفين بالتعرف على طبيعة المشكلة والنتائج المترتبة على وجودها وكيفية علاجها ويمكن لمن يرغب في معرفة المزيد الاطلاع على الكتب المتخصصة الوارد ذكرها في قائمة المراجع .

١- مشكلة عدم ثبات التباين Heteroscedasticity

من الفروض التي تقوم عليها طريقة المربعات الصغرى ، ثبات تباين الحد العشوائي ، فإذا سقط هذا الفرض ، أدى ذلك إلى أن تكون التقديرات (ب) ، (ب) ذات أخطاء معيارية كبيرة، أي تقديرات متحيزة ، ولأن هذه الأخطاء المعيارية (ع) ، (ع) تتدخل في تركيب فترات الثقة وفي اختبارات الفروض الإحصائية فهذا يعني :

- انخفاض معنوية معالم الانحدار β_1 ، β_2 .
- اتساع فترات الثقة للمعالم β_1 ، β_2 وهذا الاتساع يعد عيباً في فترة الثقة للمعالم β_1 ، β_2 .

- ضعف الثقة في نتائج اختبارات الفروض الخاصة بالمعالم β_1, β_2 .
 - ضعف الثقة في القيم المتنبأ بها للمتغير التابع بفترة ثقة .
- ويمكن الكشف عن تحقق أو عدم تحقق فرض ثبات التباين للحدود العشوائية ،
باتباع الخطوات التالية :

(اختبار جولد فيلد - كوانت) (Goldfeld & Quandt).

١- ترتب البيانات ترتيباً تصاعدياً (من الأصغر إلى الأكبر) وفق المتغير المستقل س .

٢- تحذف بعض المشاهدات الوسيطة (خمس مشاهدات مثلاً) بحيث يتبقى مجموعتين متساويتين من المشاهدات ، تحسب لكل مجموعة معادلة خط انحدار ، أى معادلة خط انحدار لقيم س الصغرى ومعادلة خط انحدار لقيم س الكبرى (إذا لم يتم حذف أية مشاهدات وسيطة يظل الاختبار صحيحاً لكن قوته في الكشف عن اختلاف التباين تكون أقل) .

٣- من معادلة خط الانحدار الأولى ، تحسب متوسط مجموع مربعات البواقي χ^2_1 ومن معادلة خط الانحدار الثانية ، تحسب متوسط مجموع مربعات البواقي χ^2_2 .

٤- أوجد النسبة ف حيث $F = \chi^2_1 / \chi^2_2$... (٢٨)

والناتج هو متغير عشوائي يتبع توزيع ف بدرجات حرية (ن-٢ - ك)
٢ ÷ حيث : ن : العدد الإجمالي للملاحظات في التجربة كلها ، و : عدد المشاهدات المحذوفة ، ك : عدد المعالم المقدرة.

٥- تقارن النسبة ف مع قيمة جدولية مستخرجة من جدول توزيع ف لنفس درجات الحرية السابقة .

٦- القرار : أ- إذا كانت ف المحسوبة أقل من ف الجدولية يقبل الفرض العدمى بأنه لا توجد مشكلة عدم ثبات التباين أى أن فرض ثبات التباين متحقق في البيانات .

ب- إذا كانت F المحسوبة أكبر من F الجدولية (أو تساويها) يرفض الفرض العدمي ويقبل الفرض البديل بأن هناك مشكلة عدم ثبات التباين أي أن فرض ثبات التباين غير متحقق في البيانات .

مثال (٧) :

يفرض أن العلاقة الانحدارية بين الأجور (ص) وعدد العاملين (س) في ٣٠ شركة كانت على الصورة التالية : $ص = ١٧,٥ + ٠,٠٩ س$. وضح كيف يتم اختبار وجود مشكلة عدم ثبات التباين على فرض توفر بيانات كاملة عن س ، ص .

الحل :

يمكن اختبار وجود مشكلة عدم ثبات التباين ، بترتيب البيانات الأصلية وفق المتغير التفسيري س ترتيباً تصاعدياً ، وإجراء انحدارين منفصلين : الأول لقيم س الصغيرة والثاني لقيم س الكبيرة ، ثم نختبر نسبة مجموع مربعات الخطأ للانحدار الثاني إلى مجموع مربعات الخطأ للانحدار الأول (أي $خ^2_2 \div خ^2_1$) . ويستخدم توزيع ف عند درجات حرية (ن - و - ٢) (ك) ٢ ÷ وهذا الاختبار ويسمى باختبار جولد فيلد - كوانت (Goldfeld & Quandt) يعد مناسباً تماماً للعينات الكبيرة (ن > ٣٠) .

بإهمال المشاهدات الست الوسطى مثلاً ونقسم الباقي إلى مجموعتين ، كل مجموعة من ١٢ مشاهدة وبحساب معادلة خط الانحدار للمجموعة الأولى ، ومعادلة خط الانحدار للمجموعة الثانية ، لنفرض أننا توصلنا إلى النتائج التالية :

$$\begin{aligned} \text{للمجموعة الأولى : } ص &= ٠,٦ + ٩ س \quad , \quad خ^2_1 = ٠,٥١ \\ \text{للمجموعة الثانية : } ص &= ٠,٢ + ١٢ س \quad , \quad خ^2_2 = ٣,٢ \end{aligned}$$

وبحساب النسبة ف حيث $ف = خ^2_2 \div خ^2_1 = ٣,٢ \div ٠,٥١ = ٦,٢٧$

وحيث أن ف الجدولية عند درجات الحرية = (ن - ١ - ٢) = (٢٠ - ١ - ٢) = ١٧
 $F_{(1,17)} = ١٠,٠٠$ (١٠, ١٠, ١٠) = ١٠,٩٧ .

وحيث أن ف المحسوبة (١٠,٩٧) تتجاوز ف الجدولية ، فإننا نرفض الفرض العدمي ونقبل البديل ، أي نقبل بوجود مشكلة عدم ثبات التباين . بالطبع هناك عدة طرق لمعالجة مشكلة عدم ثبات التباين لكن هذا الأمر خارج نطاق العرض الحالي .

٢- مشكلة الازدواج الخطي (مشكلة الارتباط بين المتغيرات المستقلة)

Multicollinearity

الارتباط بين المتغيرات المستقلة هو أحد المشاكل التي تظهر نتيجة لاختلال أحد فروض طريقة المربعات الصغرى ، وهذه المشكلة بالطبع لا توجد في حالة الانحدار البسيط لأنه يشمل متغير مستقل واحد ، ولكن توجد فقط في حالة الانحدار المتعدد أي الذي يشمل عدة متغيرات مستقلة ، وللنوضح إذا كان نموذج الانحدار المتعدد يشمل متغيرين مستقلين على الصورة : $ص = ب + س١ + س٢$ ، فإن مشكلة الازدواج الخطي تصل إلى حدها الأقصى إذا كان هناك ارتباط تام بين المتغيرات المستقلة $س١$ ، $س٢$ قيمته $+ ١$ ، وتتعلم مشكلة الازدواج الخطي إذا كان الارتباط بين المتغيرات المستقلة مساوياً للصفر ، وهنا تسمى المتغيرات المستقلة (أو التفسيرية) بالمتغيرات المتعامدة . من الناحية العملية نادراً ما يتحقق أحد الاحتمالين السابقين ، ولكن ما يحدث هو وجود درجة من الارتباط الخطي بين المتغيرات التفسيرية أكبر من الصفر وأقل من الواحد الصحيح.

وهناك العديد من الأسباب لظهور مشكلة الازدواج الخطي ، لا مجال لنذكرها الآن ويمكن الرجوع إلى الكتب المتخصصة للإلمام بها ، لكن ما يهمنا هو النتائج التي قد تترتب على وجود ازدواج خطي ، أي ارتباط بين المتغيرات التفسيرية تتراوح قيمته بين الصفر والواحد .

يترتب على وجود ازدواج خطى كبير حجم الأخطاء المعيارية للتقديرات (ب، ١، ٢) أى تصبح تقديرات متحيزة ، ومن شأن ذلك أن يؤثر على فترات الثقة وعلى اختبارات الفروض الخاصة بالمعالم β_1, β_2 ، إذ تتسع فترات الثقة لتقديرات المعالم ، وهذا يعد عيباً أو قصوراً فى فترات الثقة ، كذلك تقل درجة الثقة فى القرارات التى تتخذ بشأن معنوية المعالم β_1, β_2 . وهناك العديد من الأساليب الإحصائية التى تستخدم للكشف عن وجود الازدواج الخطى منها:

١- اختبار كلاين Klein Test

٢- اختبار الارتباط الجزئى Partial Correlation Test

٣- اختبار فارار-جلوبر Farrar – Glabuer Test

ويمكن مراجعة الكتب المتخصصة للإلمام بها .

٣- مشكلة الارتباط الذاتى للبوابات Autocorrelation

يقصد بالارتباط الذاتى وجود ارتباط بين الحدود العشوائية المتتالية X_t عبر فترات زمنية متتالية ، ووجود هذا الارتباط الذاتى يخل بأحد الفروض الأساسية التى تقوم عليها طريقة المربعات الصغرى وهذا الارتباط الذاتى غالباً ما يظهر عند تحليل بيانات السلاسل الزمنية Times series . وقياس الارتباط الذاتى لا يختلف عن قياس الارتباط العادى الذى يتم بين متغيرين ، إلا أن الارتباط الذاتى يقيس الارتباط بين القيم المتتالية لنفس المتغير X_t وهو على الصورة :

$$r = \frac{\text{مجم } X_{t-1} \text{ و } X_t}{\sqrt{\text{مجم } X_{t-1}^2 \text{ و } X_t^2}} \quad \dots (٢٩)$$

حيث X_t : الحد العشوائى فى الفترة الزمنية t .

X_{t-1} : الحد العشوائى فى الفترة الزمنية التى تسبق و مباشرة .

r : معامل الارتباط الذاتى .

ومبدئياً يمكن التعرف على وجود ارتباط ذاتي بين حدود البواقي ، برصد قيم χ مع الزمن بيانياً ، فإذا أخذت هذه البواقي شكلاً منتظماً كالدورات ، لكان ذلك دليلاً على وجود مشكلة ارتباط ذاتي بين حدود البواقي أي حدود الخطأ العشوائي .

وهناك العديد من الأسباب التي تؤدي إلى حدوث الارتباط الذاتي منها : اهمال بعض المتغيرات المستقلة وعدم إدراجها ضمن نموذج الانحدار ، افتراض صيغة رياضية خاطئة للعلاقة بين المتغيرات ، فإذا كانت العلاقة الحقيقية بين المتغيرات هي علاقة منحنى ، واستخدم الباحث علاقة خطية ، فمن شأن هذا ظهور الارتباط الذاتي .. الخ غير ذلك من الأسباب . لكن ما يهمنا هو كيفية التعرف على وجود أو عدم وجود ارتباط ذاتي . هناك العديد من الاختبارات الإحصائية يمكن الاستعانة بها للكشف عن هذه المشكلة منها :

١- اختبار فون - نيومان Von Neumann Test

٢- اختبار ديرين - واطسون Durbin - Watson Test

والاختبار الثاني من أهم وأكثر الاختبارات شيوعاً بين الإحصائيين ، ويشترط لاستخدامه أن يكون حجم العينة ١٥ أو أكثر .

اختبار ديرين - واطسون

يتم استخدام هذا الاختبار وفق الخطوات التالية :

١- الفرض العدمي : $p = 0$ صفر لا يوجد ارتباط ذاتي بين البواقي في مجتمع الدراسة ، p : معامل الارتباط الذاتي في المجتمع .

٢- الفرض البديل : $p \neq 0$ صفر

٣- احصاء ديرين - واطسون (د) على الصورة

$$D = \frac{\sum_{i=1}^{n-1} (e_i - e_{i+1})^2}{\sum_{i=1}^n e_i^2} \quad \dots (3.0)$$

ارتباط ذاتي	قرار غير حاسم	لا يوجد ارتباط ذاتي	قرار غير حاسم	ارتباط ذاتي
موجب		ذاتي ↓		سالب
صفر	١ د	٢	٣-٤	١-٤

بالتطبيع وكما ذكرنا من قبل ، فإن وجود ارتباط ذاتي بين حدود البوابي من شأنه أن تصبح الأخطاء المعيارية لتقديرات المربعات الصغرى : $E(\beta)$ ، $E(\beta)$ كبيرة أي متحيزة . وحيث أن تلك الأخطاء المعيارية تتدخل في تركيب فترات الثقة وفي اختبارات الفروض للمعامل β_1 ، فإن ذلك من شأنه أن تصبح فترات الثقة غير دقيقة نظراً لإتساع حديها ، كما أن درجة الثقة في نتائج اختبارات الفروض لا يعول عليها بدرجة كبيرة .

مثال (٨)

بفرض أنك حصلت على المبيعات التالية عن إحدى الظواهر خلال الفترة من ١٩٨٥ إلى ٢٠٠٠ ، وبفرض أن هناك متغير مستقل واحد فقط .
والمطلوب دراسة ما إذا كان هناك ارتباط ذاتي بين البواقي أم لا عند مستوى معنوية ٥%

الحل :

السنة	القيمة الفعلية ص	القيمة التقديرية ص ^٨	خ و ص - ص ^٨	خ ^٢ و	خ و - خ و ^١	خ و - خ و ^٢
١٩٨٥	٨١	٨٢	١-	١	١	١
٨٦	٨٢	٨٥	٣-	٩	٢-	٤
٨٧	٩٥	٩٢	٣+	٩	٦	٣٦
٨٨	٧٤	٧٤	.	.	٣-	٩
٨٩	٨٣	٨٢	١+	١	١	١
٩٠	٩٥	٩٧	٢-	٤	٣-	٩
٩١	٩٧	٩٥	٢+	٤	٤	١٦
٩٢	١٠٢	١٠٥	٣-	٩	٥-	٢٥
٩٣	١٠٦	١٠٨	٢-	٤	١	١
٩٤	٩٩	١٠١	٢-	٤	.	.
٩٥	١٠٠	٩٨	٢+	٤	٤	١٦
٩٦	١٠٢	١٠١	١+	١	١-	١
٩٧	١٠٨	١٠٦	٢+	٤	١	١
٩٨	١١٠	١١٢	٢-	٤	٤-	١٦
٩٩	١١٥	١١٠	٥+	٢٥	٧	٤٩
٢٠٠٠	١١٥	١١٦	١-	١	٦-	٣٦
				٨٤		٢٢٠

لدراسة الارتباط الذاتي للبواقي ، علينا بحساب عمود البواقي خ و ، حيث
خ و = ص - ص^٨ = القيمة الفعلية - القيمة التقديرية ، ثم حساب الفرق بين

خ و ١- أي الفرق بين كل قيمة والقيمة السابقة لها مباشرة وهذا ما توضحه الأعمدة في الجدول السابق

وعن طريق إحصاء ديرين - واطسون الموضح بالمعادلة المعادلة (٣٠) :

$$D = \frac{\text{مج} (X - X_{-1})^2}{\text{مج} X^2} = \frac{220}{84} = 2,62$$

خطوات الاختبار :

١- الفرض العنمي : $p = \text{صفر}$

٢- الفرض البديل : $p \neq \text{صفر}$

٣- د (إحصاء ديرين - واطسون) = ٢,٦٢

٤- القيمة الجدولية عند $\alpha = 5\%$ ، حجم العينة $n = 16$ ، وعدد المتغيرات

التفسيرية = ١ نجد أن $d = 1,1$ ، $d = 1,37$ ، ويتوقع تلك القيم على الخط

الذي يوضح حدود إحصاء ديرين واطسون

ارتباط ذاتي	قرار غير حاسم	لا يوجد ارتباط ذاتي	قرار غير حاسم	ارتباط ذاتي
ارتباط ذاتي سالب				ارتباط ذاتي موجب
٢,٩	٢,٦٣	٢	١,٣٧ = د	١,١ = د
صفر				صفر

٦- المقارنة : بتوقع قيمة $D = 2,62$ على الشكل السابق نجدها تقع في المنطقة

: لا يوجد ارتباط ذاتي بين حدود البواقي . (لا حظ أن ٢,٦٢ قريبة جدا

من ٢,٦٣ وهي منطقة قرار غير حاسم ، لذا يفضل في هذه الحالة تكبير

حجم العينة) .

مثال (٩) :

(أ) بفرض أنه في نموذج الانحدار البسيط ولعينة من ٢٠ مشاهدة ، كانت

معادلة خط الانحدار على الصورة التالية : $\hat{Y} = 56 + 0,13X$ س

وكانت قيمة إحصاء ديرين واطسون هي : $D = 0,65$ هل هناك دليل على

وجود ارتباط ذاتي بين البواقي ؟ ($\alpha = 5\%$)

(ب) بفرض أن معادلة خط الانحدار المتعدد لعينة من ٢٠ مشاهدة على الصورة التالية . ص = ٧,١٩ + ٠,٢٤ س_١ - ٠,٩٩ س_٢ وكانت قيمة احصاء ديرين - واطسون هي د = ١,٥٤ ، فهل هناك دليل على وجود ارتباط ذاتي بين البواقي ؟ ($\alpha = ٥\%$)

الحل :

(أ) حيث أن د = ٠,٦٥ أقل من د_١ = ١,٢ عند مستوى معنوية ٥% مع ن = ٢٠ ، ك = ١ ، فهناك دليل على وجود ارتباط ذاتي موجب .
(ب) حيث أن د = ١,٥٤ أكبر من د_٢ = ١,٥٣ عند مستوى معنوية ٥% مع عينة ن = ٢٠ ، ك = ٢ ، فليس هناك دليل على وجود ارتباط ذاتي .
(لاحظ أن القيم د_١ ، د_٢ تم الحصول عليها من جدول ديرين - واطسون الموضح فى نهاية الكتاب وذلك عند حجم العينة ن ، وعدد المتغيرات التفسيرية ك ومستوى المعنوية α) .

مثال (١٠) نصف محلول

بفرض أن (ص) تمثل درجات مستوى الأداء ، س_١ : المعدل التراكمي في مرحلة البكالوريوس ، س_٢ : عدد سنوات الخبرة في العمل لعينة من ٢٠ عامل .

الحل :

١- معادلة خط الانحدار المتعدد :

بالتعويض في المعادلات الطبيعية الثلاث ، نصل إلى :

$$١٣١ = ٢٠ ب + ١٤٧ ب١ + ١١٠ ب٢$$

$$١٠١٢ = ١٤٧ ب + ١٢٦١ ب١ + ٨٧٣ ب٢$$

$$٧٩٧ = ١١٠ ب + ٨٧٣ ب١ + ٧٨٨ ب٢$$

وبحل تلك المعادلات بطريقة الحذف أو المحددات أو المصفوفات ، نصل إلى :

$$ب = ٣,٤٩ ، ب١ = ٠,١٤ ، ب٢ = ٠,٣٦٨٥ = ٠,٣٧$$

∴ معادلة خط الانحدار هي : ص = ٣,٤٩ + ٠,١٤ س_١ + ٠,٣٧ س_٢

م	ص	س ١	س ٢	ص ١	س ٢	ص ٢	س ١	س ٢	ص ١	س ٢	ص ٢
١	٥	٣	١	٢٥	٩	١	١٥	٥	٣	١	٣
٢	٤	٢	٣	١٦	٤	٩	٨	١٢	٦	٦	٦
٣	٤	٤	٢	١٦	١٦	٤	١٦	٨	٨	٨	٨
٤	٩	١٢	٨								
٥	٨	١١	٧								
٦	٩	٨	٤								
٧	٧	٩	١٠								
٨	٨	٧	٥								
٩	٥	٦	٦								
١٠	٦	٥	٣								
١١	٨	٤	٩								
١٢	٤	٨	٤								
١٣	٧	٣	٧								
١٤	٦	١٢	٦								
١٥	٨	٩	٨								
١٦	٥	٨	١								
١٧	١٠	١١	١١								
١٨	٧	٧	٩								
١٩	٦	٨	٥	٣٦	٦٤	٢٥	٤٨	٣٠	٤٠		
٢٠	٥	١٠	١	٢٥	١٠٠	١	٥٠	٥	١٠		
	١٣١	١٤٧	١١٠	٩٢١	١٢٦١	٧٨٨	١٠١٢	٧٩٧	٨٧٣		

(٢) الخطأ المعياري للتقدير $\chi^2 = \frac{(\text{مجم} - \text{ص} - \text{ص}^2) \div (3 - \text{ن})}{2}$

او خمس (مج ۱ ص ۱ - مج ۱ ص ۱ - مج ۲ ص ۲ ص)

$$1,28 = 1,6332 \sqrt{\frac{27,764}{17}} =$$

(٣) معامل التحديد المتعدد (R^2) وباستخدام المعادلة (٩) نصل إلى : $R^2 = ٠,٥٥٧١$ أما معامل الارتباط المتعدد = $٠,٧٤٦$ في حين أن معامل التحديد المعدل R^2 فهو :

$$\frac{1-n}{k-n} \times (r-1) - 1 = r/2$$

$$1,00 = 1,490 - 1 = \frac{19}{17} \times (1,0071 - 1) - 1 =$$

(٤) معاملات الارتباط الجزئية :

معاملات الارتباط الجزئية $r_{2.1}$ ، $r_{3.12}$ وفق المعادلات (١٨)،

(١٩) ، تقتضي معرفة معاملات الارتباط البسيطة r_{12} ، r_{13} ، r_{23}

وهي على النحو التالي :

ر ص ۱ = ۰,۴۶۱ ، ر ص ۲ = ۰,۷۱۳ ، ر ص ۳ = ۰,۳۵۵ بالتعويض

في (١٨) ، (١٩)

$$0,6615 = 1,2 \text{ مر، } 0,3159 = 2,1 \text{ مر}$$

(٥) الخطأ المعياري لمعاملات الانحدار الجزئية : ع (ب_١) ، ع (ب_٢) :-

$$ع' (ب) = خ' مر \times$$

(مجس^۱ - (مجس^۲ - ۱) / ۱

$$[مجس^1 - (مجس^1/ن)] [مجس^2 - (مجس^2/ن)] - [مجس^2 - (مجس^2/ن)] [مجس^1 - (مجس^1/ن)] \times مجس^2/ن$$

$$\therefore 0.102 = \frac{183}{29246.4} \times 1.633 = \frac{183}{416.20 - 183 \times 182.00} \times 1.633 =$$

∴ الخطأ المعياري ع (ب) = 0.101

$$\text{بالمثل ع}^2 \text{ (ب) } = 1,633 \times \frac{182,55}{29246,4} = 0,0102$$

∴ الخطأ المعياري ع (ب) = 0,101

(٦) احسب فترة الثقة ٩٥% لمعاملات الانحدار الجزئية :

$$\beta = 0,101 \times 2,11 \pm 0,14$$

$$\beta = 0,101 \times 2,11 \pm 0,3685$$

(٧) لختبر فرض العدم القائل بأن معامل انحدار المجتمع β يساوي الصفر عند

$\alpha = 0,05$ ثم نفذ الاختبار على المعامل β باستخدام توزيع ت

$$\text{ت} = \frac{\beta - \text{ب}}{\text{ع (ب)}} = \frac{0,14 - \text{صفر}}{0,101} = 1,386$$

$$\text{ت} = \frac{\beta - \text{ب}}{\text{ع (ب)}} = \frac{0,3685 - \text{صفر}}{0,101} = 3,649$$

يلاحظ أن ت، أقل من ت الجدولية (٢,١١) ومن ثم يقبل الفرض العدمي بينما

ت، أكبر من ت الجدولية ومن ثم يقبل الفرض البديل .

(٨) اختبار العلاقة الإجمالية باستخدام تحليل التباين

$$\text{م.م.ك} = 62,95$$

$$\text{م.م. بسبب انحدار ص/س} = 0,14 \times 49,15 + 0,37 \times 76,05 = 35,186$$

جدول تحليل التباين

المصدر	م.م.	د.ح	م.م.م	نسبة التباين
انحدار ص/س	35,186	٢	17,093	10,77
د.م.م	27,764	17	1,6332	
م.م.ك	62,95	19		

وحيث أن ف الجدولية : ف (٢,١٧ ، ٥%) = ٣,٥٩ يكون القرار قبول

الفرض البديل $\beta = \beta$ صفر

(٩) اختبار مدى أهمية وجود S_1 في معادلة خط الانحدار

$$M.S. \text{ انحدار } S_1 / S_2 =$$

$$[\text{مج ص } S_2 - \text{مج ص } S_1 \times \text{مج ص } S_2 / \text{ن}] \frac{1}{183} = \frac{1}{183} (76,0) = 32 - 31,9795 =$$

جدول اختبار مغنوية S_1

المصدر	م.م	د.ح	م.م.م	نسبة التباين	ف الجدولية
م.م انحدار S_1 / S_2	32	1	3,186	1,951	ف(1, 17)
مساهمة S_1	3,186	1			4,15 =
م.م انحدار S_1 / S_2	35,186	2			
م.م.د	27,764	17	1,6332		
م.م.ك	62,95	19			

وحيث أن F المحسوبة أقل من F الجدولية ، يقبل الفرض العدمي أي أن S_1 متغير غير هام ولا ينصح ببقائه في معادلة خط الانحدار المتعدد .

(١٠) اختبار مدى أهمية وجود S_2 في معادلة خط الانحدار

$$M.S. \text{ انحدار } S_2 / S_3 = \frac{1}{180,55} (49,15) = 13,38$$

ويتكوّن جدول تحليل التباين ، نجد أن F المحسوبة = 13,38 وهي أقل من F الجدولية (4,15) مما يعني قبول الفرض البديل بأن $\beta_2 \neq 0$ ، صفر أي أنه من الضروري بقاء S_2 في معادلة خط الانحدار .

مثال (١١) : نصف محلول

الجدول التالي يبين توزيع الوزن (بالرطل) ص ، الطول (بالبوصة)

S_1 ، العمر (بالسنوات) S_2 ، لعينة من ١٢ طائب :

ص	٦٤	٧١	٥٣	٦٧	٥٥	٥٨	٧٧	٥٧	٥٦	٥١	٧٦	٦٨
س١	٥٧	٥٩	٤٩	٦٢	٥١	٥٠	٥٥	٤٨	٥٢	٤٢	٦١	٥٧
س٢	٨	١٠	٦	١١	٨	٧	١٠	٩	١٠	٦	١٢	٩

المطلوب :

- ١- تقدير معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى
- ٢- اختبار العلاقة الإجمالية بين المتغيرات الثلاث مستخدماً أسلوب تحليل التباين عند مستوى معنوية ٥% .
- ٣- معامل التحديد ومعامل التحديد المعدل ومعامل الارتباط المتعدد .
- ٤- اختبار مدى أهمية وجود العمر (س٢) في معادلة خط الانحدار ثم كرر التحليل بالنسبة للطول (س١) عن مستوى معنوية ١% .

الحل :

(١) من بيانات الجدول ، يمكن أن نصل إلى المجاميع التالية .

$$\begin{aligned} \text{مجم ص} &= ٧٥٣ \quad \text{مجم س١} = ٦٤٣ \quad \text{مجم س٢} = ١٠٦ \\ \text{مجم ص}^2 &= ٤٨١٣٩ \quad \text{مجم س١}^2 = ٣٤٨٤٣ \quad \text{مجم س٢}^2 = ٩٧٦ \\ \text{مجم ص س١} &= ٤٠٨٣٠ \quad \text{مجم ص س٢} = ٦٧٩٦ \quad \text{مجم س١ س٢} = ٥٧٧٩ \end{aligned}$$

معادلة خط الانحدار المتعدد :

$$\text{ص} = ٣,٦٥١٢ + ٠,٨٥٤٦ \text{ س١} + ١,٥٠٦٣ \text{ س٢}$$

$$(٢) \text{ م.م.ك} = ٤٨١٣٩ - ١٢/٢(٧٥٣) = ٨٨٨,٢٥$$

$$\text{م.م.ب} = ٤٨١,٧٥ \times ٠,٨٥٤٦ + ١,٥٠٦٣ \times ١٤٤,٥ = ٦٢٩,٣٦$$

$$\text{ف المحسوبة} = \frac{٣١٤,٦٨}{٢٨,٧٦} = ١٠,٩٤$$

$$\text{ف الجدولية : ف (٢ ، ٩ ، ١) = ٨,٠٢}$$

$$(٣) \text{ معامل التحديد ر}^2 = \frac{٦٢٩,٣٦}{٨٨٨,٢٥} = ٠,٧٠٨٥ \quad \text{ومعامل الارتباط} = ٠,٨٤١٧$$

$$\text{أما معامل التحديد المعدل } R^2 = 1 - \frac{(1 - R^2) \times (n - 1)}{n - k}$$

$$= 1 - \frac{11}{9} \times 0,2915 = 0,644$$

(٤) اختبار مدى أهمية S_2 (العمر) أي اختبار $\beta_2 = 0$ - صفر

$$F = \frac{(481,75)}{388,917} = 1,238$$

بعد تكوين جدول تحليل التباين ، نجد أن : مساهمة $S_2 = 32,618$ ونجد أيضاً أن : ف المحسوبة = $1,134$ وهي أقل من ف (١ ، ٩ ، ٥%) = $5,12$. وعلى ذلك نعتبر S_2 غير معنوية ، أي أن وجود العمر على حدة في معادلة خط الانحدار لا يفيد معنوياً في تفسير التغير في الوزن ، وبالتالي لا يفيد معنوياً في التنبؤ بقيم الوزن في غيبة الطول .

وعند اختبار مدى أهمية طول القامة (S_1) ، أي اختبار $\beta_1 = 0$ - صفر نجد أن :

$$F = \frac{(114,5)}{39,667} = 2,888$$

وبعد تكوين جدول تحليل التباين ، نجد أن مساهمة $S_1 = 102,972$ وسنجد أن ف المحسوبة لها = $3,58$ وهي أقل من ف الجدولية (٥,١٢) مما يعني قبول الفرض العدمي أي أن S_1 يعتبر متغير غير معنوي في غيبة المتغير S_2 بمعنى أنه ليس من المفيد استخدام S_1 على حدة في غيبة S_2 لتفسير التغير في الوزن ص . مما تقدم نجد أن التغير في الوزن يفسر في ضوء وجود كلا المتغيرين معا S_1 ، S_2 وأن أحدهما لا يغني عن الآخر في تفسير التغير في الوزن .

مثال (١٣) : نصف محلول

الجدول التالي يعطي دخل الفرد الحقيقي بالآلاف دولار (ص) ، نسبة القوة العاملة في الزراعة (S_1) ، متوسط سنوات التعليم (S_2) لعدد ١٥ دولة

متقدمة في عام ١٩٨١ . المطلوب :

- ١- معادلة خط الانحدار المتعدد بطريقة المربعات الصغرى .
- ٢- الخطأ المعياري لمعادلة خط الانحدار .
- ٣- تبين التقديرات ب، ب_١، والأخطاء المعيارية لها .
- ٤- اختبار معنوية المعالم β ، β_1 ، عند مستوى معنوية ٥% باستخدام توزيع ت وكذلك فترات الثقة لتلك المعالم .
- ٥- معامل التحديد ومعامل التحديد المعدل .
- ٦- اختبار المعنوية الإجمالية لمعادلة خط الانحدار باستخدام تحليل التباين .
- ٧- معاملات الارتباط الجزئية . والبيانات كما يلي :

ص	٦	٨	٨	٧	٧	١٢	٩	٨	٩	١٠	١١	٩	١٠	١١
س١	٩	١٠	٨	٧	١٠	٤	٥	٥	٥	٦	٧	٤	٩	٥
س٢	٨	١٣	١١	١٠	١٢	١٦	١٠	١٠	١٢	١٤	١٢	١٦	١٤	١٠

الحل :

١- معادلة خط الانحدار المتعدد :

$$\text{ص} = ٦,٢٦ - ٠,٣٨ \text{ س} + ٠,٤٥ \text{ س}٢$$

٢- تبين البواقي (تبين تقديرات معادلة خط الانحدار) خ^٢ من

$$\text{خ}^٢ \text{ من} = \frac{\text{مج}(\text{ص} - \text{ص}^٨)}{\text{ن} - ٣} = \frac{١٢,٢٧٣}{١٢} = ١,٠٢$$

∴ الخطأ المعياري خمر = ١,٠٠٩٩ = ١ تقريباً

٣- تبين التقديرات ب ، ب_١

$$\text{ع}^٢(\text{ب}) = ٠,٠٢ ، \text{ع}^٢(\text{ب}١) = ٠,٠١$$

لما الأخطاء المعيارية ع(ب) = ٠,١٤ ، ع(ب_١) = ٠,١

٤- اختبار معنوية المعالم β ، β_1 باستخدام توزيع ت

$$\text{ت}١ = \frac{-٠,٣٨}{٠,١٤} = -٢,٧١ ، \text{ت}٢ = \frac{٠,٤٥}{٠,١} = ٤,٥$$

أما ت الجدولية : ت (١٢ ، ٠.٠٢٥) = ٢,١٧٩

∴ β_١ معنوية ، β_٢ معنوية

أما فترات الثقة للمعالم β_١ ، β_٢ :

$$\beta_1 = -0.38 - 0.14 \times 2.179 \pm 0.07 = -0.69$$

$$\beta_2 = 0.45 - 0.1 \times 2.179 \pm 0.23 = 0.67$$

٥- معامل التحديد ومعامل التحديد المعدل :

$$R^2 = 0.6932$$

$$R^2_{\text{المعدل}} = 0.641$$

٦- اختبار المعنوية الإجمالية لمعادلة خط الانحدار :

$$F_{\text{المحسوبة}} = \frac{27.727}{12 \div 12.273} = 13.09$$

وهي أكبر من ف الجدولية (٣,٨٨) عند مستوى معنوية ٥% ودرجات حرية ٢,١٢

ويمكن حساب ف بدلالة R^٢ :

$$F = \frac{R^2}{1-R^2} \times \frac{n-k}{k-1} = \frac{0.6932}{0.6932-1} \times \frac{12}{2} = 13.06$$

٧- معاملات الارتباط الجزئية :

هنا يلزم معرفة معاملات الارتباط البسيطة وهي :

$$r_{11} = 0.5715$$

$$r_{21} = 0.6984$$

$$r_{22} = 0.8072$$

تمارين

- ١- أ- ما هي فروض نموذج الانحدار المتعدد ؟
 ب- ما هو المقصود بمعامل الانحدار الجزئي ومعامل الارتباط الجزئي؟
 ج- ما هو المقصود بمعامل التحديد ومعامل التحديد المعدل ؟
 د- إذا كانت قيمة معامل التحديد في نموذج الانحدار البسيط هي ٠,٧ ، فهل تكون قيمة معامل التحديد في نموذج الانحدار المتعدد أكبر أم أقل من ٠,٧ ؟ ولماذا ؟
 هـ- ما هي أهمية الخطأ المعياري لتقدير معاملات خط الانحدار ؟

٢- ترغب إحدى شركات التأمين على الحياة في تحديد العلاقة بين حجم مبيعات مندوبيها وبين خبرتهم في هذا المجال وحجم أسرة كل منهم . سحبت عينة عشوائية من تسعة مندوبين وسجلت لهم : المبيعات (ص) بالآلاف جنيه ، الخبرة (س) بالسنوات ، عدد أفراد الأسرة (س٢) والنتائج هي :

ص	١	٢	٣	٤	٥	٦	٧	٨	٩
س١	٢	١	٣	٣	٤	٥	٦	٥	٧
س٢	١	٢	١	٣	١	٤	٧	٥	٦

المطلوب :

- أ- معادلة خط انحدار ص/س١ ، س١ ، مستخدماً طريقة المربعات الصغرى .
- ب- قدر حجم المبيعات السنوية لمندوب خبرته ١٠ سنوات وأسرته من ٥ أفراد .
- ج- حدد الخطأ المعياري للتقدير .
- د- احسب معاملات الارتباط الجزئية .

- هـ- أوجد فترة الثقة ٩٥% لمعاملات الانحدار β_1, β_2 .
- و- اختبار فرض العدم القائل بأن كل معامل انحدار جزئي يساوى الصفر فى مقابل أنه لا يساوى الصفر عند مستوى معنوية ٥% .
- ز - اختبار العلاقة الإجمالية لمعادلة خط الانحدار مستخدماً أسلوب تحليل التباين .
- ح- اختبر مدى أهمية وجود المتغير س_١ فى معادلة خط الانحدار .

٣- أجريت دراسة لتحديد العلاقة بين درجات مستوى الأداء (ص) لعشرة مهندسين . وبين الخبرة (س_١) وعمر كل واحد منهم (س_٢) وكانت كما يلي:

ص	١	٢	٣	٤	٥	٦	٧	٨	٩
س _١	١	٣	٢	٥	٥	٤	٧	٩	٨
س _٢	٢٤	٢٤	٣٥	٣٢	٣١	٣٨	٣٦	٥٠	٤٢

المطلوب :

- أ- معادلة خط الانحدار المتعدد ص/س_١، س_٢ .
- ب- معامل التحديد ومعامل التحديد المعدل .
- ج- اختبار معنوية β عند $\alpha = ٥\%$ مستخدماً أسلوب فترة الثقة ثم دعم ذلك مستخدماً توزيع ت .
- د - اختبار $\beta_1 = \beta_2 = ٠$ صفر باستخدام تحليل التباين ، $\alpha = ٥\%$.
- ٤- البيانات التالية تمثل عدد مرات الإجازات المرضية (ص) وعدد سنوات الخدمة (س_١) وعدد أفراد الأسرة (س_٢) لعينة عشوائية من خمس موظفين بإحدى الشركات :

ص	١٠	١٦	١٤	١٥	١٥
س _١	١٥	٩	١٣	١١	١٢
س _٢	٦	٢	٤	٥	٣

المطلوب :

- أ- معادلة خط انحدار ص/س، س، ص.
- ب- عدد الإجازات المرضية لموظف بالشركة خدمته ١٥ سنة وعدد أفراد أسرته أربعة أفراد.
- ج- احسب الخطأ المعياري لتقدير معادلة خط الانحدار.
- د- احسب معامل التحديد ومعامل التحديد المعدل ومنلول كل منهما.
- هـ- احسب فترة الثقة ٩٥% لمعامل الانحدار β .
- و- اختبر معنوية المعامل β عند $\alpha = 5\%$. مستخدماً توزيع ت ثم دعم ذلك باستخدام أسلوب تحليل التباين.
- ز- اختبر معنوية المتغير س، عند $\alpha = 5\%$ بطريقتين مختلفتين.
- ح- أجريت دراسة عن العلاقة بين درجة التحصيل في مادة الإحصاء (ص) وبين كل من عدد ساعات المذاكرة (س) والعمر (س) على عينة من الدارسين وكانت النتائج كما يلي :

الطالب	أ	ب	ج	د	هـ	و
ص	٣٨	٦٧	٤٩	١٢	١٠	٢٤
س	٩	١١	١٤	٤	٦	٦
س	٢٠	١٩	٢١	٢٦	٢٤	٣٠

المطلوب :-

- أ- معادلة خط أنحد ص / س، س، ص.
- ب- اختبار β - صفر هند $\alpha = 5\%$ مستخدماً أسلوب فترة الثقة للعملة β ، ثم دعم قرارك باستخدام أسلوب اختبار المعنوية (اختبار ت) ثم ناقش علاقة هذا الاختبار باختبار مدى أهمية المتغير س، في معادلة خط الانحدار.
- ج- معامل التحديد ومعامل التحديد المعدل ثم قارن بينهما.
- د- معاملات الارتباط الجزئية.

٦- البيانات التالية تمثل كمية فول الصويا (ص) وكمية المياه المستخدمة في الري (س)، وكمية السماد (س_٢) لعينة عشوائية من ست أفدنة .

ص	٢٠	٢٨	٣٢	٢٥	٣٠	٣١
س	٢	٣	٤	١	٣	٥
س _٢	٢	٣	٣	٢	٢	٣

المطلوب :

- أ- معادلة خط انحدار ص/س، س_٢
- ب- فترة الثقة ٩٥% لكل من معاملي الانحدار الجزئي في المجتمع .
- ج- اختبار العلاقة الإجمالية لمعادلة خط الانحدار باستخدام أسلوب تحليل التباين .
- هـ- اختبار مدى أهمية وجود س_٢ في معادلة خط الانحدار عند $\alpha = 5\%$ ، ثم كرر التحليل بالنسبة للتغير س_٢ .

٧- بفرض أنك حصلت على البيانات التالية :

ص	٢٠	٢٨	٤٠	٤٥	٣٧	٥٢	٥٤	٤٣	٦٥	٥٦
س	٢	٣	٥	٤	٣	٥	٧	٦	٧	٨
س _٢	٥	٦	٦	٥	٥	٧	٦	٦	٧	٧

المطلوب : تأكد من النتائج التالية :

- أ- معادلة خط الانحدار المتعدد : ص = ٤,٧٦ + ٥,٢٩ س + ٢,١٣ س_٢ .
- ب- الخط المعياري للتقدير = ٧,٠٧١ ، الخطأ المعياري لمعاملات خط الانحدار : ع(ب) = ١,٧٨ ، ع(ب) = ١٨,٩٥ .
- ج- β ، معنوية إحصائيا عند $\alpha = 5\%$.
- د- β ، ليست معنوية إحصائيا عند $\alpha = 5\%$.
- د- فترة الثقة ٩٥% لمعامل خط الانحدار هي : $\beta = ١,٠٨ ، ٩,٥$ & $\beta = -٨,١٦ ، -١٢,٤٢$

- هـ- معامل التحديد $R^2 = 0.79$ ، معامل التحديد المعدل $R^2 = 0.73$.
 و- مجموع المربعات بسبب إندثار ص/ص_١ = ٦٤٩ .
 مجموع مربعات البواقي = ٥٠ ، ف المحسوبة (٢،٧) = ١٢،٩٨ ويقبل للفرض البديل β صفر .
 ز- معاملات الارتباط الجزئية : $R_{١١} = 0.74$ ، $R_{١٢} = 0.18$.
 أى متغير مستقل يساهم أكثر في القدرة التفسيرية للنموذج ؟
 الإجابة : ص_١ .

- ٨- الجدول التالي يعطى بيانات عينة عشوائية من ١٢ أسرة ، تشمل عدد الأطفال في الأسرة (ص) ، عدد الأطفال الذين كانوا يرغبون في إنجابهم وقت الزواج (ص_١) ، وعدد سنوات التعليم للزوجة (ص_٢) .

الأسرة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
ص	٤	٣	٠	٤	٤	٣	٠	٤	٣	١	٣	١
ص _١	٣	٣	٠	٢	٢	٣	٠	٣	٢	١	٣	٢
ص _٢	١٢	١٤	١٨	١٠	١٠	١٤	١٨	١٢	١٥	١٦	١٤	١٥

المطلوب :

- أ- معادلة خط الإندثار المتعدد ص / ص_١ ، ص_٢ .
 ب- إختبار معنوية β ، β عند $\alpha = 5\%$ مستخدماً توزيع ت.
 ج- أوجد معامل التحديد و معامل التحديد المعدل ثم قارن بينهما .
 د- إختبار المعنوية الأجمالية لمعادلة خط الإندثار عند $\alpha = 5\%$.
 هـ- أوجد معاملات الارتباط الجزئية وحدد أى المتغيرات المستقلة يساهم أكثر في القدرة التفسيرية للنموذج .

ملحوظة : تأكد من الإجابات التالية :

أ- ص = $6.9 + 0.53 \text{ ص}_1 - 0.39 \text{ ص}_2$.

ب- حيث أن $t = 3,12$ و $t = -5,57$ ، فإن كلا من β_1 ، β_2 ،

تصبح معنوية إحصائياً عند $\alpha = 5\%$.

ج- $r^2 = 0,92$ ، $r^2 = 0,90$.

د- $F(2, 9) = 51,31$.

هـ- $r_{12} = 0,71$ و $r_{13} = 0,87$ ، بالتالي فإن r_1 تساهم أكثر

من r_2 في القدرة التفسيرية للنموذج .

٩- تبين من أحد التقارير أن : (ص) تمثل حجم المبيعات ، (س) عدد ساعات

العمل ، (س) عدد أشهر الخبرة وذلك لعينة عشوائية مكونة من ١٠ عمال

من إحدى الشركات :

مج ص = ١٠٣ ، مج س = ١٠٢ ، مج ص س = ٧٥ =

مج ص^٢ = ١٣٦٩ ، مج ص س = ٢٨٨ ، مج ص س = ٩٨٩ =

وبفرض أنك حصلت على معادلة خط الانحدار المتعدد على الصورة .

ص = -٢,٥ + ١,١٠٣ س_١ + ٠,٢٠٧ س_٢

المطلوب :

اختبار العلاقة الإجمالية لمعادلة خط الانحدار عند $\alpha = 5\%$.

١٠- (أ) بفرض أن هناك ٥٠ مشاهدة وأربع متغيرات تفسيرية ، ما هو

تعليقك حول الارتباط الذاتي إذا كانت قيمة إحصاء ديدين -

واطمس على الصور التالية : $(\alpha = 5\%)$

د = ١,٠٥ ، د = ١,٤ ، د = ٢,٥ ، د = ٣,٩٧

(ب) بفرض أنك حصلت على النموذجين التاليين ، كل منها على عينة

من ١٦ مفردة

ص = ٠,٤٥ - ٠,٠٠٤ س ، د = ٠,٨٢٥

ص = ٠,٤٧ - ٠,٠١ س_١ + ٠,٠٠٥ س_٢ ، د = ١,٢٨

حدد أي النموذجين به ارتباط ذاتي عند $\alpha = 5\%$

الفصل الخامس
الطرق الالاعلمية
Non-Parametric Methods

محتويات الفصل:

- (١) : مقدمة.
- (٢) : توزيع كا^٢.
- (٣) : إختبار الإشارة.
- (٤) : إختبار مجموع الرتب / إختبار U: مان-هوتيني.
- (٥) : إختبار ولكوكسن التي تعتمد على الرتب.
- (٦) : تحليل التباين في اتجاه واحد بالرتب / إختبار كروسكال-والس.
- (٧) : معامل إرتباط الرتب لسبيرمان وإختبار معنويته.

تمارين

الفصل الخامس

الطرق الالاعملية

Non-Parametric Methods

(١) مقدمة :

سبق أن عرفنا الإستنتاج الرياضى بأنه الوصول إلى قرارات بشأن المتغير موضوع الدراسة فى المجتمع إعتياداً على نتائج تسجل عن مفردات عشوائية تسحب من ذلك المجتمع ولقد تعرضنا - فى هذا المجال - إلى بعض الإختبارات الإحصائية المتعلقة بمتوسط واحد أو الفرق بين متوسطين ثم التى تعتمد على عدة متوسطات وكانت أدوات الإختبارات الإحصائية عبارة عن متغيرات تربط ما بين القيمة المحسوبة من عينة عشوائية واحدة أو أكثر والقيمة الفرضية للمتوسط (أو المتوسطات) - تحت شرط صحة الفرض العدمى - بتوزيع احتمالى معين يمكن تعيين احتمالاته من جداول خاصة بكل توزيع احتمالى. ولقد كانت هذه الإختبارات تفترض - لصلاحية استخدامها - تحقق فروضاً معينة تتصل بنوع القياسات وأنها قياسات كمية لفترة (Interval) أو نسبية (Ratio) وأن توزيع المتغير فى المجتمع يتبع التوزيع المعتاد الطبيعى فضلاً عن شروطاً أخرى تتعلق بتساوى التباينات داخل المعالجات وأن الخطأ العشوائى مستقل فى توزيعه عن المعالجات وأنها تتبع فى توزيعها التوزيع المعتاد.

ومع ذلك فإننا كثيراً ما نواجه مجالات لا نستطيع فيها أن نجزم بشكل التوزيع وأن أحسن ما يمكن أن يفترض بشأنه أن توزيع مستمر كما أن القياسات قد تكون نوعية (Norminal) أو ترتيبية (Ordinal) فمثلاً لو أن ترتيب مفردات عينة عشوائية من الوحدات الصناعية بحسب درجة الأمن الصناعى ومعدل إصابات العمل كانت على النحو التالى (جدول ٥-١):

جدول (١)

التوزيع الترتيبى لعينة عشوائية من الوحدات الصناعية بحسب درجة الأمن الصناعى ومعدل

إصابات العمل								
الوحدة الصناعية	١	٢	٣	٤	٥	٦	٧	٨
الترتيب بحسب درجة الأمن الصناعى:	٦	٣	٤	١	٧	٢	٨	٥
الترتيب بحسب معدل إصابات العمل:	٥	١	٨	٢	٦	٣	٧	٤

فإن بيانات كهذه لا تعطى تفصيلات كافية عن درجة الأمن الصناعى ومعدل إصابات العمل ولا تعنى بالضرورة أن الوحدة الصناعية رقم (٣) يصل معدل إصابات العمل بها ثمانية أمثال الوحدة رقم (٢) ولو أن البيانات أعطيت فى صورة مستوى أو درجة الأمن الصناعى ومعدل إصابة العمل لأمكن قياس العلاقة بين المتغيرين بالإرتباط أو الإنحدار ولأمكن استخدام أسلوب تحليل الإنحدار لإختيار معنوية العلاقة بين درجة الأمن الصناعى وإصابة العمل. أما وأن القياسات ترتيبية على هذا النحو فلا بد من وسيلة أخرى ثلاث طبيعة القياسات كما أنه فى حالات أخرى فإنه بدافع السهولة وإمكانية التصحية ببعض الدقة فى سبيل تسهيل العمليات الحسابية وسرعة الوصول إلى قرار فإننا فى أى من الحالات الثلاث (عدم معرفة شكل التوزيع وطبيعة القياسات وأنها ترتيبية أو نوعية (أو وصفية) والسهولة والسرعة) فإننا نلجأ إلى أساليب إحصائية لا تعتمد على معالم أو مؤشرات أو بارامترات تعرف باسم الأساليب الإحصائية الالعملية أو اللابارامترية أو التى لا تعتمد على توزيع معين (Free or Non-parametric Distribution).

وهذه الأساليب وإن كانت مفيدة فى الحالات التى ذكرناها إلا أنه يعاب عليها أنها تتجاهل قدرأ من المعلومات التى قد تكون متاحة عن المتغير إذ أن استخدام الترتيب قد يتغاضى عن الفرق فى القيمة بين مفردتين متتاليتين فى الترتيب مما يجعلها أقل حساسية للقيم القليلة المتطرفة ولكنها لهذا السبب عادة ما

تكون أقل كفاءة من الأساليب البارامترية. كما أنه في الحالات التي يمكن فيها استخدام اختبار بارامترى (اختبارات مثلاً) وآخر لا بارامترى (اختبار الإشارة مثلاً) فإن الأخير عادة ما يكون أقل قوة من سابقه وهو ما سوف نشير إليه في الفقرة التالية الخاصة بالقوة المكافئة كما سنشير إليه في حينه عندما تتوافر شروط استخدام المدخلين في الاختبار.

وهذه الأساليب قد تهدف إلى الوصف دون التحليل والإستنتاج ومنها الوسيط والمتوال ونصف المدى الربيعي ومعامل الإنواء الربيعي ومعامل ارتباط الرتب ومعامل التوافق وهو ما عرضنا له في مرحلة سابقة. كما أنها قد تهدف إلى الإستنتاج الإحصائي وإتخاذ القرارات وهو ما يعنينا في المرحلة الحالية. وسوف نعرض نماذج من الإختبارات الإحصائية اللابارامترية أو الالامعلمية وفي جميع الأحوال فإن خطوات الإختبار الإحصائي سوف تكون واحدة مهما اختلفت أداة الإختبار بارامترى أو لابارامترى.

القوة المكافئة للإختبار:

تعرف قوة الإختبار بأنها احتمال رفض الفرض العدمى وهو صحيح أى قدرة الإختبار على تجنب الخطأ من النوع الثانى أو حماية الباحث من الوقوع فى هذا الخطأ الذى يشار إليه بأنه خطأ من النوع الثانى.

وعادة ما يستخدم هذا المعيار للمقارنة بين الإختبارات الإحصائية اللابارامترية ونظائرها البارامترية. ولهذا الغرض سوف نعرف قاعدة القوة المكافئة للإختبار (Power-efficiency) على النحو التالى:

إذا كان حجم العينة اللازم لإجراء الإختبار أ هو n_1 وهو الحجم الذى يحقق للإختبار قوة معينة وكان يمكن الوصول إلى ذات الدرجة من القوة بإختبار آخر ب بعينة حجمها n_2 فإن:

$$\text{القوة المكافئة للإختبار ب} = (n_1 + n_2) \times 100$$

فمثلاً إذا كان لإجراء الاختبار أ بقوة معينة فإنه يلزم سحب عينة عشوائية حجمها $n = 20$ مفردة. وللوصول إلى ذات القوة باختبار آخر ب فإنه يلزم سحب عينة عشوائية حجمها $n = 25$ مفردة فإن:

$$\frac{20}{25} = 0.8 = 80\%$$

بمعنى أننا نحتاج إلى 100 مفردة باستخدام الاختبار ب لتحقيق نفس الدرجة من القوة باستخدام 80 وحدة فقط مع استخدام الاختبار أ.

وسوف نتعرض فى بقية هذا الباب لأمتلة من الاختبارات الإحصائية الالابارامترية:

⊗ باستخدام عينة واحدة.

⊗ باستخدام عينتين مستقلتين.

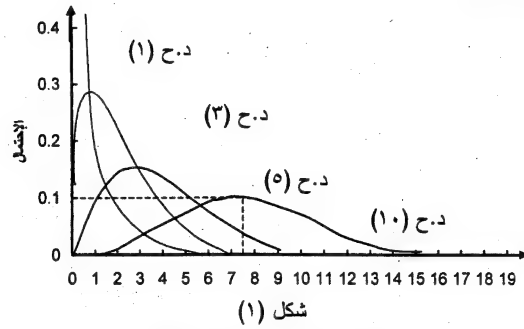
⊗ باستخدام عدة عينات مستقلة.

مع الإشارة إلى الاختبار الالابارامترى المناظر-إذا ما توافرت شروط استخدامه-وقو الاختبار الالابارامترى مقارناً بقوة الاختبار الالابارامترى المناظر. وسوف نبدأ بتقديم توزيع احتمالى شائع الاستخدام فى الاختبارات الالاعلمية خاصة فى حالة البيانات التصنيفية وهو توزيع كا².

(٢) توزيع كا² -Distribution χ^2 :

توزيع كا² هو توزيع احتمالى اكتشفه الإحصائيان المعروفان سير رونالد فيشر وكارل بيرسون R.A. Fisher & Karl Person فى أوائل القرن الحالى ولايستوقف شكل التوزيع على عدد مفردات الدراسة ولكن يتحدد شكل التوزيع على أساس عدد درجات الحرية وبالتالى فهو ليس توزيعاً وحيداً ولكنها مجموعة من توزيعات تختلف باختلاف درجات الحرية. والشكل التالى يوضح صوراً مختلفة لهذا التوزيع لأعداد مختلفة من درجات الحرية ويبين منه أن توزيع كا² يكون شديد الإنواء إلى اليمين متى كان عدد درجات الحرية صغيراً ولكنه يأخذ

فى التماثل كلما زاد عدد درجات الحرية أنظر شكل (١-٥) التالى - ويعطى رقم () الإحتمالات أو المساحات تحت منحنى توزيع كاي^٢ Chi-Square بدرجات حرية مختلفة.



شكل (١) توزيع كاي^٢ لدرجات حرية مختلفة

وأنه وإن كان هذا التوزيع قد عرف باستخدامه فى الإختبارات الإحصائية حين تكون القراءات نوعية أو تصنيفية (Categorical) أو حين تتعدد النسب حيث يتعذر استخدام إختبارات النسبة التى تعرضنا لها فى فصل سابق إلا أنه يمكن استخدامه مع بقية أنواع القياسات الترتيبية أو بفترة أو نسبية متى أمكن رصد هذه المشاهدات عن المتغير مقابل فئات.

ويستخدم هذا التوزيع فى إجراء الإختبارات التالية:

- ١- اختبار جودة المطابقة أى مدى تشابه مشاهد لتوزيع نظرى وهو اختبار يعتمد على عينة واحدة.
- ٢- اختبار استقلال ظاهرتين وهو اختبار يعتمد أيضاً على عينة واحدة.
- ٣- اختبار تجانس توزيع عدة ظواهر فى مجتمع واحد وهو يعتمد على عدة عينات.

٤- الاختبارات الخاصة بالتباين σ^2 وتقديرها بفترة ثقة.
ولكننا سوف نقتصر فى المرحلة الحالية على اختبار جودة المطابقة واستقلال توزيع ظاهرتين فى مجتمع واحد واختبار تجانس توزيع ظاهرة ما فى عدة مجتمعات. مع إشارة عابرة إلى استخدام χ^2 فى الاختبارات المتعلقة بالتباين σ^2 وتقديرها بفترة ثقة.

أ: اختبار χ^2 لجودة المطابقة :

كثيراً ما يعنى الباحث بدراسة ما إذا كان التوزيع المشاهد لظاهرة ما فى أحد المجتمعات لا يختلف عن توزيع ظاهرى أو متوقع لهذه الظاهرة حيث يستحدد هذا التوزيع النظرى على أساس فروض معينة ، وإجراء اختبار كهذا فإنه يمكن مقارنة التكرارات المشاهدة (H_r) للظاهرة موضوع الدراسة بالتكرارات النظرية أو المتوقعة (T_r) ، $r = 1, 2, \dots$ ، ك تمثل الخلايا التى تتوزع داخلها التكرارات المشاهدة والنظرية وكلما اقترب التوزيع المشاهد من التوزيع النظرى كلما قلت قيمة الفروق $H_r - T_r$ وتتعلم الفروق تماماً عندما يتطابق التوزيعين المشاهد والنظرى وبالتالي فإنه يمكن أن تكون هذه الفروق أساساً لوسيلة أو أداة اختبار إحصائى ويمكن أن نثبت أن المتغير العشوائى:

$$\chi^2 = \sum \frac{(H_r - T_r)^2}{T_r} \quad (1)$$

له توزيع χ^2 ب درجات حرية ك - ١

ويمكن استخدام قيم التوزيع الاحتمالى χ^2 ب درجات حرية ك-١ فى إجراء اختبار كهذا حيث ك : عدد الخلايا أو الفئات أو النواتج الممكنة للتجربة. ويمكن تلخيص خطوات إجراء اختبار جودة المطابقة فى الأتى:

(١) نبدأ بتحديد الفرض المراد اختباره أى تحديد التوزيع النظرى للمتغير وتقدير القيم المتوقعة فى العينة العشوائية على أساس هذا الفرض النظرى ويرصد

مقابلها القراءات المشاهدة وترمز للكرارات المشاهدة فى كل فئة أو خلية بالرمز h_r وللكرارات المتوقعة t_r أى أنه سيكون لدينا r من الكرارات المشاهدة h_r بناظرها r من الكرارات المتوقعة t_r .

$$(2) \text{ نحسب المقدار: } \chi^2 = \sum_r \frac{(h_r - t_r)^2}{t_r} \text{ والجمع بالنسبة لجميع}$$

الخلايا كـ.

(3) نحدد مستوى المعنوية α منذ البداية (5% أو 1%) أو أى قيمة أخرى يختارها الباحث.

(4) ونقارن قيمة χ^2 (المحسوبة من البيانات) فى الخطوة السابقة بقيمة

$\chi^2_{(k-1, \alpha)}$ من الجدول الخاص بتوزيع χ^2 الموضح فى جدول (د)

بالملاحق. ونرفض الفرض الأصلى بعدم وجود اختلاف معنوى بين

التوزيع المشاهد والتوزيع النظرى عند مستوى المعنوية α إذا كانت χ^2

المحسوبة لقيمة χ^2 من الجدول ويقبل الفرض العدمى فيما عدا ذلك. أى أن

الاختبار فى هذه الحالة هو اختبار فى اتجاه واحد طرف أيمن.

وتوضح الأمثلة التالية طريقة إجراء هذا الاختبار.

مثال (1):

تعد جداول الأعداد العشوائية بحيث تكون الأرقام من صفر إلى 9 فى

ترتيب عشوائى فى كل عدد ، وبحيث تكون فرص ظهور هذه الأرقام

متساوية وتساوى 0.1 بالنسبة لكل رقم.

وللتحقق من توافر هذه الخاصية فى تركيب جداول الأعداد العشوائية

(جدول رقم 7) بالملاحق ، فقد أخذت مفردات العمودين الأول والثانى من

الصفحة الثانية من الجدول وكانت نتيجة حصر الكرارات المشاهدة لكل رقم

وكذلك الكرارات المقدمة لكل كما يوضحها الجدول التالى جنول (5-2). اختبر

فرض جودة المطابقة - أى مطابقة تركيب الأعداد العشوائية لخاصية إعداد تلك الجداول عند مستوى المعنوية ٥٪.

الحل :

جدول (٢)

التوزيع المشاهد والنظري (المتوقع)

للأرقام بالأعداد العشوائية

الرقم	التكرار المشاهد (هـ)	التكرار المتوقع (ت ر)	كا
صفر	٢٦	٢٥	٠,٠٤
١	٢١	٢٥	٠,٠٤
٢	٢٥	٢٥	صفر
٣	٢٥	٢٥	صفر
٤	٢٨	٢٥	٠,٣٦
٥	٢٧	٢٥	٠,١٦
٦	٢١	٢٥	٠,٦٤
٧	٢٩	٢٥	٠,٦٤
٨	٢٣	٢٥	٠,١٦
٩	٢٥	٢٥	صفر
مجموع	٢٥٠	٢٥٠	٢,٦٤

ويتم اختبار فرض جودة المطابقة على النحو التالى:

(١) الفرض : أن جداول الأعداد العشوائية - الجزء الذى تم

استخدامه فى الاختبار - تتوافر له خاصية الجداول

العشوائية.

(٢) مستوى المعنوية : $\alpha = ٥\%$

$$(3) \text{ أداة الاختبار : } \chi^2 = \frac{(n-1) s^2}{\sigma^2}$$

يرفض الفرض العدمى (وهو الفرض بعدم اختلاف التوزيع المشاهد عن التوزيع النظرى أو المتوقع) إذا كانت $\chi^2 \leq \chi^2_{(1-\alpha, n-1)}$ ، $\alpha = 5\%$ من الجدول = ١٦,٩٠٩ .

$$(4) \text{ ومن البيانات : } \chi^2 = 2,64 < 16,919$$

∴ يقبل فرض العدم أى أن جداول الأعداد العشوائية يتفق فى تركيبه مع خاصية الجداول العشوائية.

مثال (٣) :

الجدول التالى يوضح التوزيع التكرارى لنتائج اختبار ما طبق على عينة عشوائية مكونة من ٨٦ عاملاً فى نهاية برنامج تدريبى. اختبر الفرض بأن توزيع درجات الاختبار تتبع توزيعاً معتاداً متوسطه وتباينه هو نفس متوسط وتباين توزيع العينة وذلك عند مستوى المعنوية ٥% .

ملحوظة : لم تحدد قيم المؤشرات التى تعين التوزيع المعتمد النظرى فى هذا المثال ولذلك تعين علينا تقديرها من بيانات العينة وهذه المؤشرات هى $\mu = \bar{x} = 26,1$ درجة ، $\sigma = s = 6,45$ درجة.

الحل :

جدول (٣)

التوزيع النظري والمُشاهد للدرجات في برنامج تدريبي لعند ٨٦ عاملاً

الفئات	التكرارات المُشاهدة H_i	الحدود العليا s	$s - \bar{x} =$ ع	الاحتمالات التجميعية	التكرارات ت	كأ
٧-٥	صفر	٧,٥	٢,٨٨-	٠,٠٠٢	٠,٢	
١٠-٨	صفر	١٠,٥	٢,٤٢-	٠,٠٠٨	٠,٢	
١٣-١١	٤	١٣,٥	١,٩٥-	٠,٠١٦	١,٥	٠,٢٤٨
١٦-١٤	٣	١٦,٥	١,٤٩-	٠,٠٦٨	٣,٦	
١٩-١٧	٨	١٩,٥	١,٠٢-	٠,١٥٤	٧,٤	٠,٤٩
٢٢-٢٠	١٣	٢٢,٥	٠,٥٦-	٠,٢٨٨	١١,٦	٠,١٦٩
٢٥-٢٣	٩	٢٥,٥	٠,٠٩-	٠,٤٦٤	١٥,١	٢,٤٦٤
٢٨-٢٦	١٧	٢٨,٥	٠,٣٧	٠,٨٠٠	١٥,٥	٠,١٤٥
٣١-٢٩	١	٣١,٥	٠,٨٤	٠,٩٠٣	١٣,٤	٠,٠٢٧
٣٤-٣٢	٨	٣٤,٥	١,٣٠	٠,٩٦٢	٨,٩	٠,٠٩١
٣٧-٣٥	٥	٣٧,٥	١,٧٧	٠,٩٨٧	٥,٠	
٤٠-٣٨	٤	٤٠,٥	٢,٢٣	٠,٩٩٧	٣,٢	
٤٣-٤١	١	٤٣,٥	٢,٧٠	٠,٩٩٩	٠,٨	٠,٣٤٨
٤٦-٤٤	صفر	٤٦,٥	٢,١٦	١,٠٠٠	٠,٢	
< ٤٧	صفر					
	٨٦					٣,٥٤١

وبمقارنة $\chi^2 = ٣,٥٤١$ بقيمة $\chi^2_{٨-١-٢} = ٠,٠٥$ من

الجدول $= ١١,٠٧٠$ نجد أن $\chi^2 < \chi^2_{٨-١-٢}$ من الجدول وبالتالي يقبل الفرض

الأصلي بأن التوزيع التكراري المُشاهد يتبع توزيعاً معتمداً متوسطه

$\mu = ٢٦,١$ وتباينه $= ٦,٤٥$ ويلاحظ في هذا المثال أن:

(١) التكرارات المتوقعة ت ر فى الفئات الأربعة الأولى وفى كل من الخمسة الأخيرة تقل عن ٥ تكرارات فى كل فئة لذلك فقد أدمجت الفئات الأربع الأولى وكذلك الخمس الأخيرة وجمعت تكراراتها المتوقعة والملاحظة لتحقيق التقارب لتوزيع كا^٢ وهذه قاعدة عملية شائعة الاستخدام ويتعين استخدامها حينما نقل التكرارات المتوقعة فى خلية أو أكثر عن ٥ وحدات ثم تحسب قيمة كا^٢ بعد إجراء هذا الإدماج أى أن كا^٢ فى الفئات الأربع الأولى (١٦-٥)

$$٠,٢٤٨ = \frac{\chi^2(٥,٨ - ٧)}{٥,٨}$$

$$وبالمثل فإن كا^٢ فى الخمس الأخيرة ٣٥ فأكثر = \frac{\chi^2(٨,٣ - ١٠)}{٨,٣} = ٠,٣٤٨$$

(٢) إن المتغير العشوائى $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X - \mu}{\sigma}$ له توزيع معتاد

معيارى (ى) متوسطه الصفر وتباينه الوحدة وبالتالي فإنه يمكن إيجاد المساحة المحصورة بين -∞ وأى قيمة (ى) على المحور الأفقى باستخدام جداول التوزيع المعتاد المعيارى.

فمثلاً :

$$ح (-\infty \text{ إلى } ٢,٨٢) = ٠,٠٠٢$$

$$ح (-\infty \text{ إلى } ٢,٤٢) = ٠,٠٠٨$$

$$ح (-\infty \text{ إلى } ١,٩٥) = ٠,٠١٦$$

وتكون المساحة المناظرة للفئة ٥ - ٧ = احتمال وقوع مشاهدة فى هذه الفترة = ٠,٠٠٢

$$\therefore \text{التكرار المتوقع} = \frac{٢ \times ٨٦}{١٠٠٠} = ٠,٢$$

والمساحة المناظرة للفئة ٨ - ١٠ = احتمال وقوع مشاهدة فى هذه الفترة = ٠,٠٠٨ - ٠,٠٠٢ = ٠,٠٠٦

ويكون التكرار المتوقع = $\frac{2 \times 86}{1000} = 0,0$ وهكذا.

(٣) عدد درجات الحرية لقيمة كا^٢ = عدد الفئات (أو الخلايا) - ١ - عدد المؤشرات أو المعالم التي حسبت من العينة. وفي مثالنا هذا درجات الحرية = عدد الخلايا بعد إدماج الأربعة الأولى والخمس الأخيرة - ١ - درجتين مقابل تقدير المتوسط الحقيقي والتباين الحقيقي = ٨ - ١ - ٢ = ٥.

ملحوظة:

(١) إذا أجرى اختبار يعتمد على توزيع كا^٢ بدرجة حرية واحدة أى كانت القراءات المشاهدة والمتوقعة تقع فى خليتين اثنتين فقط فإنه يتعين أن تكون القراءات المتوقعة فى كل خلية لا تقل عن خمسة. وإذا كان عدد الخلايا ك < ٢ فإنه يتعين ألا يزيد عدد الخلايا التى تقل فيها القراءات المتوقعة عن ٥ عن ٢٠٪ من مجموع عدد الخلايا أو كان التكرار المتوقع فى إحداها يقل عن ١ فإنه يمكن إدماج الخلايا المتعاقبة كالمثال (٥-٢) وغنى عن الذكر أن كا^٢ دائماً موجبة القيمة أما إذا كانت قيمة كا^٢ = صفر فإن ذلك يستوجب إعادة فحص البيانات وكيف جمعت فقد تكون الطريقة التى صنفنا البيانات على أساسها قد أدت إلى عدم ظهور فروقاً حقيقية بين القراءات المشاهدة والمتوقعة.

استخدام توزيع كا^٢ لاختبار جودة المطابقة لتوزيع غير مستمر

متعدد النواتج:

كما يمكن استخدام توزيع كا^٢ فى اختبار التجارب العشوائية متعددة النواتج أى ك ، ك ، ك < ٢ (multinomial). فتجربة إلقاء زهرة النرد ينتج عنها ٦ نواتج متماثلة مستقلة وثابتة عند تكرار التجربة (النواتج هى ١، ٢، ٣، ٤، ٥، ٦). وبالمثل فإن استطلاع رأى فى موضوع ما يختلف من الموافقة إلى الرفض إلى الامتناع عن إبداء رأى وهى أيضاً نواتج متماثلة مستقلة وثابتة وفى جميع

الأحوال فإن $L_1 + L_2 + L_3 + \dots + L_k = 1$. وفى حالات كهذه فإن الاختبارات التى تعتمد على توزيع ذو الحدين أو الطبيعى المعيارى لا تصلح للاستخدام والبديل هنا هو توزيع كاي^٢ وهو ما يوضحه المثال التالى:

مثال (٣):

أجريت دراسة تسويقية للتعرف على مدى تفضيل المستهلكين لسلعة ما (المسظفات الصناعية مثلاً) من أنواع مختلفة عددها ٥ أنواع فقد سحبت عينة عشوائية مكونة من ١٠٠٠ من المشتريين لهذه السلعة وكان عدد من فضلوا نوعاً على الأنواع الأخرى كالتالى:

النوع	أ	ب	ج	د	هـ	المجموع
عدد الذين فضلوا هذا النوع على غيره	٢١٤	٢٣١	١٨٢	١٥٤	٢١٩	١٠٠٠

عند مستوى المعنوية ١٪ اختبر ما إذا كانت نسبة من فضلوا نوعاً على الآخر لا تختلف من نوع لآخر.

الحل:

(١) الفرض العدمى هو أن $L_1 = L_2 = L_3 = \dots = L_k$.

$$\text{حيث } L_i = \frac{n_i}{n} = 0.20$$

الفرض البديل هو : أن احتمال التفضيل لنوع على الآخر $\neq 0.20$ لنوعين من الأنواع الخمسة على الأقل.

(٢) أداة الاختبار هى:

$$\text{كا}^2 = \sum \frac{(n_i - E_i)^2}{E_i} \text{ لها توزيع :}$$

$$\text{كا}^2 \text{ بدرجات حرية } (k - 1) = (5 - 1) = 4$$

$$\text{ويرفض الفرض العدمى إذا كانت } \text{كا}^2 < 13.277$$

(٣) ومن البيانات :

$$\begin{aligned} \chi^2_{\text{كا}} &= \frac{(200 - 182)^2}{200} + \frac{(200 - 231)^2}{200} + \frac{(200 - 214)^2}{200} \\ &+ \frac{(200 - 219)^2}{200} + \frac{(200 - 154)^2}{200} \\ &= 1,800 + 18,080 + 1,620 + 4,800 + 0,980 \\ &= 19,790 \end{aligned}$$

(٤) القرار : $\chi^2_{\text{كا}} = 19,790 < 13,277$

∴ يرفض الفرض العدمي ويقبل الفرض البديل.

وهذا يلاحظ أن الجزء الأكبر من اختلاف $\chi^2_{\text{كا}}$ عن $\chi^2_{\text{يرجع}}$ يرجع إلى اختلاف

معنوي في نسبة الذين فضلوا النوع ب ، د عن بقية الأنواع.

$$\text{وهذه النسبة} = \frac{(10,580 - 4,800)}{200} \times 100 = 77,74\%$$

ويفيد مثل هذا التحليل في اتخاذ قرار تسويقي مناسب بالنسبة لمبيعات هذه

الأنواع الخمسة.

ب : اختبار χ^2 لاستقلال توزيع ظاهرتين أو متغيرين في مجتمع واحد :

يهدف هذا الاختبار إلى اختبار استقلال توزيع متغيرين في مجتمع واحد ،

كالمستوى التعليمي لرب الأسرة ودخله السنوي أو حجم الأسرة أو عدد الأطفال

والمستوى التعليمي للأب أو للأم.

وقد تكون قياسات أحد المتغيرين أو كلاهما وصفية.

وتختار عينة عشوائية واحدة حجمها ن ثم تصنف مفردات العينة بحسب

فئات كل من المتغيرين في جداول توافق مكون من ل صف ، م عمود.

خطوات الاختبار :

(١) يجرى اختبار توزيع المتغيرين بتقدير التكرارات المتوقعة في كل خلية

من خلايا جدول التوافق (Contingency table) وعددها ل × م خلية.

وإذا رمزنا إلى التكرارات المشاهدة في كل خلية بالرمز h_{rj} والتكرارات المتوقعة بالرمز t_{rj} حيث r عدد الصفوف ، و عدد الأعمدة ، ($r = 1, 2, \dots, l$ ، و $j = 1, 2, \dots, m$) وحيث تتحدد t_{rj} في حالة صحة فرض استقلال المتغيرين كالآتي:

$$t_{rj} = \left(\frac{\sum_{j=1}^m h_{rj}}{\sum_{r=1}^l \sum_{j=1}^m h_{rj}} \right) \times \left(\sum_{r=1}^l h_{rj} \right) \quad (2)$$

= مجموع الصف الذي تقع للقراءات المشاهدة \times مجموع العمود الذي تقع فيه هذه القراءة المجموع الكلي

وذلك تطبيقاً لقاعدة ضرب الاحتمالات في حالة الاستقلال.

(2) وإذا صح فرض الاستقلال فإن المتغير العشوائى:

$$\chi^2 = \sum_{r=1}^l \sum_{j=1}^m \frac{(h_{rj} - t_{rj})^2}{t_{rj}}$$

له توزيع كاً بدرجات حرية ($l-1$) ($m-1$) (3)

(3) تحسب كاً من بيانات العينة وتقارن بقيمة كاً ($\alpha, 1-m, 1-l$) ويقبل أو يرفض فرض الاستقلال حسب الأحوال.

والمثال التالى يوضح هذا الأسلوب:

مثال (4):

سحبت عينة عشوائية مكونة من 400 شخصاً ثم صنفوا إجاباتهم عن المستوى التعليمى لكل منهم وجملة دخله السنوى فكانت النتائج كما يوضحها الجدول التالى :

اختبر الفرض باستقلال الدخل السنوى عن المستوى التعليمى عند مستوى المعنوية 5% .

مجموع الصف الأول × مجموع العمود الأول

$$ن \times ح(١١) ح(١١) = ن \times \frac{\text{مجموع الكلي}}{٤٠٠} = \frac{١٢٠ \times ١٧٠}{٤٠٠} = \frac{١٢٠}{٤٠٠} \times \frac{١٧٠}{٤٠٠} = ٥١ \text{ شخصاً}$$

وبالمثل بالنسبة لباقي الخلايا وتكون التكرارات المتوقعة لجميع خلايا الجدول كما هو الحال بالجدول (٤) السابق.
ويستكمل الاختبار كالآتي:

(١) الفرض الأصلي : إن الدخل السنوي مستقل عن المستوى التعليمي والفرض البديل إنهما غير مستقلين.

(٢) مستوى المعنوية $\alpha = ٥\%$

(٣) وسيلة الاختبار $\chi^2 = \frac{\text{مجموع (هـ - ت)}}{\text{ت}}$

والجمع بالنسبة لجميع الخلايا في جميع الصفوف وجميع الأعمدة. ويرفض الفرض الأصلي إذا كانت أي المحسوبة χ^2 (ل-م) $(١-\alpha)$.

$$\begin{aligned} (٤) \text{ ومن البيانات فإن : } \chi^2 &= \frac{(٥١-٦٣)^2}{٥١} + \frac{(٥٢,٥-٤٨)^2}{٥٢,٥} + \frac{(١٦,٥-٩)^2}{١٦,٥} \\ &+ \frac{(٧٦,٥-٧٨)^2}{٧٨,٥} + \frac{(٧٨,٧٥-٨٢)^2}{٧٨,٧٥} + \frac{(٢٤,٧٥-٢٠)^2}{٢٤,٧٥} \\ &+ \frac{(٤٢,٥-٢٩)^2}{٤٢,٥} + \frac{(٤٣,٧٥-٤٥)^2}{٤٣,٧٥} + \frac{(١٣,٧٥-٢٦)^2}{١٣,٧٥} \end{aligned}$$

$$= ٢,٢٨ + ٠,٣٨ + ٣,٤١ + ٠,٠٣ + ٠,٩٢ =$$

$$= ٤,٢٩ + ٠,٢٤ + ١٠,٩١ = ٢٥,٤٤$$

ولما كانت χ^2 $٩,٤٨٨ > ٢٥,٤٤$ لذلك نرفض الفرض باستقلال توزيع الدخل السنوي عن المستوى التعليمي لصاحب هذا الدخل عند مستوى المعنوية ٥% .

ملعوظة:

يمكن من ملاحظة القيم الللى تكون فى مجموعها كا^٢ المحسوبة ومعرفة أيهما أكبر مساهمة فى تكوين هذا المجموع وبالتالي يمكن معرفة أى فئات الدخل أكثر من غيرها إعماداً على المستوى التعليمى وتطبيقاً لذلك فى المثال السابق يتضح أن الدخل المرتفع قد أسهم فى قيمة كا^٢ بـ ١٠,٩١+٠,٩٢+٢,٢٨ = ١٤,١١ أى حوالى ٦٣% مما يعنى أن ارتفاع الدخل له أثر معنوى فى تحقيق مستوى تعليمى مرتفع.

حالة خاصة: اختبار كا^٢ : جدول ٢ × ٢ :

إذا كانت البيانات مبوية فى جدول مزدوج مكون من صفين وعمودين فإنه يمكن الوصول إلى قيمة كا^٢ بالطريقة التالية بدلاً من الطريقة العامة الللى شرحناها فى الفقرات السابقة.

المجموع	الخاصية (١)		الخاصية (٢)
	الفئة (٢/١)	الفئة (١/١)	
ا + ب	ب	ا	الفئة (١/٢)
ج + د	د	ج	الفئة (٢/٢)
ن	ب + د	ا + ج	المجموع

$$\text{كا}^2 \text{ المحسوبة} = \frac{(ا د - ب ج)ن}{(ا + ج)(ب + د)(د + ج)(ب + د)} \quad (٤)$$

مثال (٥):

الجدول التالى رقم (٥) يوضح توزيع مفردات عينة عشوائية من ٤١٢ فرداً من الريف والحضر بحسب موافقتهم أو رفضهم لمشروع التأمينات الإجتماعية. اختبر عند مستوى المعنوية ٥% استقلال الرأى عن محل الإقامة.

الحل :

جدول (۵)

توزيع عينة من ٤١٢ فرداً حسب محل الإقامة الرأى فى التأمينات الإجتماعية

المجموع	الرأى		الإقامة
	معارض	موافق	
١٩٥	١١١	٨٤	ريف
٢١٧	٩٥	١٢٢	حضر
٤١٣	٢٠٦	٢٠٦	المجموع

$$V_{1.} = \frac{Y(111 \times 122 - 90 \times 185)}{190 \times 217 \times 2.7 \times 2.7} = 0.15$$

وعند مستوى المعنوية ٥٪ فإن $\chi^2_{(1-2)} (1-2)$ من الجدول = ٣,٨٤

أى أن $\text{كا}^2 = 7,10 < \text{كا}^2$ من الجدول = 3,84

لذلك يرفض الفرض بأنه لا فرق بين تقبل أفراد الريف والحضر للمشروع عند مستوى المعنوية ٥% .

وكان يمكن حل نفس المثال باستخدام الطريقة العامة أى حساب $\frac{d}{dt}$

استخدام وسيلة الاختبار كا^٢ = مج $\frac{(م - ت)^2}{ت}$ مع الجمع بالنسبة لجميع

الخلايا وكنا سنحصل على نفس القيمة وبالتالي نصل إلى نفس القرار.

تصحیح بیتس :

يستخدم تصحيح بيتس لتصحيح قيمة κ^* المحسوبة لمقارنتها بـ κ^* من الجدول وتوضح أهمية هذا التصحيح إذا ما لاحظنا أن توزيع κ^* هو توزيع لمتغير مستمر بينما المتغير الذي تحسب من بياناته κ^* متغير منقطع. ويستخدم هذا التصحيح عندما تكون درجات الحرية لـ κ^* درجة واحدة وذلك إذا كان

التوزيع مكونة من خليتين فقط تقع فيها التكرارات المشاهدة والمتوقعة وذلك في حالة إختيار جودة المطابقة ، وفي حالة مطابقة التوزيع المشاهد لتوزيع ذي حدين. كما يكون اختبار كا^٢ بدرجة حرية واحدة إذا كنا بصدد جدول 2×2 في اختبارات الاستقلال والتجانس.

وفي الحالة الأولى يتم التصحيح بطرح ٠,٥ من الفرق المطلق بين التكرار المشاهد والمتوقع في كل من الخليتين. أي أن كا^٢ في هذه الحالة تكون على الصورة.

$$\text{كا}^2 (\text{المصححة}) = \frac{(أ_١ - أ_١ت - ٠,٥)^2}{أ_١ت} + \frac{(أ_٢ - أ_٢ت - ٠,٥)^2}{أ_٢ت} \quad (٥)$$

أما في حالة جدول 2×2 فإننا نحصل على الصيغة المصححة كالآتي: ٠,٥

$$\text{كا}^2 (\text{المصححة}) = \frac{(أ - ب - ج - د - ١)^2}{(أ + ب + ج + د)(أ + ب + ج + د + ١)} \quad (٦)$$

وقد يؤثر هذا التصحيح على القرار أو رفض الفرض الأصلي إذا كانت قيمة كا^٢ المحسوبة بدون تصحيح قريبة من القيمة المعنوية التي تؤدي إلى رفض الفرض الأصلي. إذ أن التصحيح قد يؤدي إلى خفض قيمة كا^٢ المحسوبة وبالتالي إلى تغيير القرار من الرفض إلى القبول أما إذا كانت قيمة كا^٢ المحسوبة لا تؤدي إلى رفض الفرض الأصلي أو أن قيمتها تختلف كثيراً عن قيمة كا^٢ من الجدول فإن هذا التصحيح يفقد قيمته العملية.

مثال (٦):

يوضح الجدول التالي رقم (٦) نتائج دراسة أجريت على عينة مكونة من ٤٥ فرداً وتفضيلهم الأنباء المذاعة من الإذاعة على المنشورة بالصحف اختبر عند مستوى ٥٪ استقلال الرأي عن نوع الفرد.

جدول (٦)

توزيع أفراد عينة مكونة من ٤٥ فرداً بحسب النوع والرأى فى موضوع ما

المجموع	الرأى		الإجابة
	إناث	ذكور	
٣١	١٠	٢١	نعم
١٤	١٠	٤	لا
٤٥	٢٠	٢٥	المجموع

الحل :

من البيانات الموضحة بالجدول السابق يتبين أن:

$$\text{كا}^* (\text{المحسوبة}) = \frac{45^2 (10 \times 4 - 10 \times 21)}{31 \times 14 \times 20 \times 25} = 0.99$$

$$\text{كا}^* (\text{المصححة}) = \frac{45^2 \left(\frac{45}{2} - 10 \times 4 - 10 \times 21 \right)}{31 \times 14 \times 20 \times 25} = 4.01$$

وقيمة كا^{*} المحسوبة تتجاوز قيمة كا^{*} من الجدول بدرجة حرية واحدة عند $\alpha = 5\%$ وتساوى ٣,٨٤ إذا حسبت كا^{*} بدون تصحيح كما أن لتصحيح لا يؤدي إلى تغيير فى القرار ولم يكن من المحتمل أصلاً أن نقل القيمة المحسوبة بعد التصحيح عن ٣,٨٤ لذلك كان يمكن أن يصرف النظر عن التصحيح.

ج: اختبار كا^٢ لتجانس توزيع ظاهرة ما فى عدة مجتمعات :

لا تختلف طريقة إجراء اختبار كهذا عن الإختبار الخاص باستقلال ظاهرتين باستخدام توزيع كا^٢ ولكنهما يختلفان من حيث طريقة بناء النموذج وتصميم التجربة وبالتالي الفرض موضوع الاختبار. ففي حالة استقلال ظاهرتين فإن حجم العينة مثبت أو بمعنى آخر فإننا نختار عينة واحدة من

المجتمع ويتم تصنيف مفرداتها بحسب متغيرين أو صفتين وعليه فإن المجاميع الهامشية للصغوف والأعمدة غير ماثبة أما فى حالة اختبار التجانس (Homogeneity) أى عدم اختلاف توزيع المتغير فى عدة مجتمعات فيتم سحب عدة عينات من كل مجتمع عينة واحدة ذات حجم محدد ويتم رصد التوزيع المشاهد للمتغير فى كل عينة وبالتالي فإن المجاميع الهامشية ماثبة فى إتجاه واحد هو مجموع كل عينة.

خطوات الاختبار :

(١) الفرض العدمى: هو أنه لا يوجد اختلاف فى توزيع المتغير فى المجتمعات المختلفة أى أن التوزيع متجانس فى المجتمعات المختلفة أما الفرض البديل: فهو أن التوزيع يختلف من مجتمع لآخر.

(٢) يحدد مستوى المعنوية α .

كا^٢ = مجرر $\frac{(ت - ت)^2}{ت}$ حيث $ت$ تمثل القراءات المشاهدة أى توزيع كل عينة على فئات المتغير ، $ت$ تمثل التكرارات المتوقعة لكل تكرار مشاهد وتحسب كالمعتاد.

$$تدر = \frac{\text{مجموع الصف (ر) } \times \text{مجموع العمود (و)}}{ن}$$

لجميع $ر = ١, ٢, \dots, ل$ ، و $و = ١, ٢, \dots, م$ ، $ن$: الحجم الكلى للعينات.

(٤) وإذا صح الفرض العدمى فإن المتغير العشوائى:

$$كا^{**} = \text{مجرر} \frac{(ت - ت)^2}{ت} \text{ له توزيع كا}^2 \text{ بدرجات حرية (ل-١) (م-١)}$$

ويقبل أو يرفض الفرض العدمى بمقارنة كا^{**} بـ كا^٢ (ل-١) (م-١).

مثال (٧):

يبين الجدول (٧) توزيع من اجتازوا الاختبار في نهاية برنامج تدريبي موحد طبق في ثلاثة أقسام مختلفة من إحدى المؤسسات. اختبر الفرض القائل بأن قدرات المتدربين متقاربة في الأقسام الثلاثة عند مستوى المعنوية ٥% .

جدول (٧)

توزيع التقديرات في برنامج تدريبي في عينات من ثلاث أقسام مختلفة

المجموع	نتيجة الاختبار				القسم
	فشل		نجاح		
	رت٢	رت٣	رت١	رت٢	
٥٥	٨,٢٥	٥	٤٦,٧٥	٥٠	الأول
٦١	٩,١٥	١٤	٥١,٥٨	٤٧	الثاني
٦٤	٩,٦	٨	٥٤,٦٧	٥٦	الثالث
١٨٠	٢٧	٢٧	١٥٣	١٥٣	المجموع

وتحسب التكرارات المتوقعة كالآتي:

$$٤٦,٧٥ = \frac{١٥٣ \times ٥٥}{١٨٠} = ١١$$

$$٥١,٥٨ = \frac{١٥٣ \times ٦١}{١٨٠} = ١٢$$

$$٥٤,٦٧ = \frac{١٥٣ \times ٦٤}{١٨٠} = ١٣$$

$$٨,٢٥ = ٤٦,٧٥ - ٥٥ = \frac{٢٧ \times ٥٥}{١٨٠} = ٨,٢٥$$

$$٩,١٥ = ٥١,٥٨ - ٦١ = \frac{٢٧ \times ٦١}{١٨٠} = ٩,١٥$$

$$٩,٦٠ = ٥٤,٦٧ - ٦٤ = \frac{٢٧ \times ٦٤}{١٨٠} = ٩,٦٠$$

ويتم استكمال خطوات الاختبار كالآتي:

- (١) الفرض الأصلي: أنه في ضوء نتيجة الاختبار فإن قدرات المتدربين لا تختلف معنوياً أى أن قدراتهم متجانسة فى الأقسام الثلاثة.
- (٢) مستوى المعنوية ٥٪.
- (٣) وسيلة الاختبار χ^2 المحسوبة = مجرور $\frac{\chi^2 (ت - هـ)}{ت}$
- ويرفض الفرض الأصلي إذا كانت χ^2 المحسوبة $\leq \chi^2$ من الجدول عند درجات الحرية (١-ل) (١-م) ومستوى المعنوية ٥٪.
- (٤) ومن البيانات يتضح أن:

$$\chi^2_{\text{كا}} = \frac{\chi^2 (٥٤,٦٧-٥٦)}{٥٤,٦٧} + \frac{\chi^2 (٥١,٨٥-٤٧)}{٥١,٨٥} + \frac{\chi^2 (٤٦,٧٥-٥٠)}{٤٦,٧٥} + \frac{\chi^2 (٩,٦-٨)}{٩,٦} + \frac{\chi^2 (٩,١٥-١٤)}{٩,١٥} + \frac{\chi^2 (٨,٢٥-٥)}{٨,٢٥}$$

وهى أصغر من $\chi^2_{\text{كا}} = ٥,٩٩$ (٠,٠٥, ٧)

لذلك يقبل الفرض العدمى بعدم وجود اختلاف معنوى فى توزيع قدرات المتدربين فى الأقسام الثلاث عن مستوى المعنوية ٥٪.

ومن البديهي أننا وقد قبلنا الفرض العدمى فلا محل لتحليل مكونات χ^2 لتحديد أى الأقسام أسهمت أكثر من غيرها فى تكوين المجموع لـ χ^2 . ولو أننا رفضنا الفرض بتجانس توزيع الدرجات فى المجتمعات الثلاث لكان علينا محاولة تفسير درجة إختلاف المجتمعات الثلاث بالرجوع إلى قيمة الحدود المختلفة التى تكون فى مجموعها القيمة الكلية لـ χ^2 .

ملاحظات ختامية بشأن استخدام χ^2 فى الاختبارات الالاعلمية:

- (١) إن استخدام χ^2 فى الاختبارات الإحصائية الخاصة بجودة المطابقة أو استقلال توزيع ظاهرتين لا يفترض شكلاً معيناً لتوزيع الظاهرة أو الظاهرتين فى المجتمع الأصلي ولذلك يطلق عليه اختبار التوزيع الحر

(distribution free test) بمعنى أنه لايشترط تحقق توزيع معين للظاهرة في المجتمع.

(٢) أن هذا التوزيع - مع توافر الشروط الخاصة بعدد التكرارات المتوقعة في كل خلية - يصلح للاستخدام في حالة القياسات النوعية والترتيبية والقياسات النسبية أو بفترة متى أمكن تصنيف أو تبويب القراءات للمشاهدة في فئات مختلفة.

(٣) إذا كان جدول التوافق الذي وزعت فيه مفردات العينات بحسب فئات المتغيرات هو جدول 2×2 فيمكن تطبيق أسلوب كا^٢ (2×2) الحالة الخاصة كما أوضحنا في السابق.

د: استخدام توزيع كا^٢ في الاستنتاج الإحصائي بشأن تباين المجتمع σ^2 :
يستخدم توزيع كا^٢ في الاستنتاج الإحصائي بشأن تباين المجتمع σ^2 وذلك اعتماداً على أن توزيع العينات للمتغير العشوائي.

$$\frac{(n-1)S^2}{\sigma^2} \quad (٧)$$

له توزيع كا^٢ بدرجات حرية (ن - ١) متى كانت العينة العشوائية مسحوبة من مجتمع معناد (أو قريباً من المعناد) وحيث ع^٢ التباين محسوباً من العينة العشوائية بالعلاقة مج (س - م)^٢ ÷ ن - ١ .

وتأسيساً على ذلك فإنه يمكن تقدير σ^2 بفترة ثقة $(1 - \alpha) \times 100\%$ كالآتي:

$$\sigma^2 = (n-1)S^2 \div \text{كا}^2_{\frac{\alpha}{2}} \text{ كحد أدنى.}$$

$$= (n-1)S^2 \div \text{كا}^2_{1-\frac{\alpha}{2}} \text{ كحد أعلى.}$$

كما يمكن استخدام المتغير العشوائي

$$\frac{(n-1)S^2}{\sigma^2} = \text{كا}^2$$

كأداة لاختبارات الفروض المتعلقة بـ ٢٥ (أنظر تمرين ٨ ، ٩).

* اختبار المتتابعات The Runs Test :

نحن نفترض دائماً عند إجراء اختبار إحصائي ما أو عند قياس ظاهرة ما أن القياسات مسجلة من مفردات عينة عشوائية وإن كن لا نتحقق من عشوائية العينة وبالتالي فإنه إن لم تكن العينة عشوائية فإن المقاييس الإحصائية التي تحسب قد تكون متحيزة كما أن القرارات التي تتخذ في شأن الاختبارات الإحصائية قد تكون غير صحيحة.

فإذا ألقيت قطعة عملة ٢٠ مرة في تجربة عشوائية لتسجيل عدد مرات ظهور الوجه الذي يحمل الصورة أو الكتابة وبفرض أن نتيجة التجربة كانت كالآتي:

ص ص ص ص ص ص ص ص ص ص
ك ك ك ك ك ك ك ك ك ك

أو كانت على النحو التالي:

ص ك ص ك ص ك ص ك ص ك
ص ك ص ك ص ك ص ك ص ك

فإن كلا من الناتجين يشير إلى أنهما نواتج غير عشوائية إذ أن تتابع ظهور الصورة والكتابة في الحالتين لا يتفق مع العشوائية لـ $n = 20$ حيث تشمل السلسلة الأولى على متابعتين وهو عدد قليل قد يكون نادر التحقق وتشمل الثانية على ٢٠ متتابعة وهو عدد كبير لا يتحقق إلا باحتمال صغير للغاية.

فإذا عرفت المتتابعات (Runs) بأنها ظهور قيمة أو وجه من أوجه المتغير مرة أو أكثر مسبقة أو متبوعة بظهور القيمة أو الوجه الآخر للمتغير مرة أو أكثر فإن توزيع العينات للمتتابعات يعطى مؤشراً يصلح للتعرف أو الحكم على عشوائية العينة وبالتالي أنها تمثل المجتمع الذي سحبت منه.

١: اختبار المتتابعات في حالة العينات الصغيرة $n \geq 20$:

خطوات الاختبار:

- دع n_1 = عدد العناصر من نوع ما أو وجه من أوجه المتغير ،
 n_2 = عدد العناصر من نوع ثان من المتغير أو وجهه الآخر ،
 $n = n_1 + n_2$ = حجم العينة.
 مـ : عدد المتتابعات.

فإذا كانت n_1 ، n_2 كلاهما ≥ 20 فإن ملحق رقم (٨) يعطى الحدين الأدنى والأعلى لقبول العدد المشاهد من المتتابعات والحكم بأن العينة عشوائية. وأي عدد يقل عن الحد الأدنى أو يزيد عن الحد الأعلى يعني أن احتمال مشاهدة عدد من المتتابعات يساوى أو يقل عن الحد الأدنى $\alpha = \frac{1}{p}$. كما أن احتمال مشاهدة عدد من المتتابعات مساو للحد الأعلى أو أكبر منه $\alpha = \frac{1}{p}$. وفى كلتا الحالتين يرفض الفرض بعشوائية العينة عند مستوى المعنوية α مما يترتب عليه رفض الفرض بعشوائية العينة.
 فمثلاً عند $n_1 = n_2 = 10$ وعند $\alpha = 5\%$ فإن الحد الأدنى لـ مـ ≥ 6 والحد الأعلى ≤ 16 واحتمال الحصول على مـ ≥ 6 أو مـ ≤ 16 كل يتحقق باحتمال قدره ٠.٠٢٥ . هذا ويمكن إجراء الاختبار في طرف واحد.

مثال (٨):

للتعرف على رأى قائدى السيارات فى إجراءات تنظيم المرور وأثره على إنسياب حركة المرور فقد سجل رأى عشرون من قائدى السيارات عند تقاطعين رئيسيين فى مدينة القاهرة وكانت إجابات أفراد العينة بالموافقة (نعم) أو الرفض (لا) كما هو موضح بجدول (٨) التالى:

جدول (٨)

توزيع آراء ٢٠ من قاندى السيارات بشأن تنظيم المرور

١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	مسلسل
نعم	نعم	لا	لا	نعم	لا	لا	لا	نعم	نعم	الإجابة
٥		٤		٣		٢		١		المتتابعات
٢٠	١٩	١٨	١٧	١٦	١٥	١٤	١٣	١٢	١١	مسلسل
نعم	لا	لا	لا	نعم	لا	نعم	لا	لا	لا	الإجابة
١١		١٠		٩		٨		٧		المتتابعات

هل هناك شك فى أن هذه الإجابات صدرت عن عينة عشوائية من قاندى

السيارات عند $\alpha = 5\%$ ؟

الحل :

$$n = 20, \quad n_1 = 8$$

$$m^* = 11 \quad \text{متتابعة}$$

الفرض الأسمى : أن العينة من الإجابات العشوائية.

عند $\alpha = 5\%$ فإن $m^* \geq 6$ ، $m^* \leq 16$ تؤدي إلى رفض الفرض

بعشوائية العينة.

ولكن $m^* = 11$ لذلك يقبل الفرض عند $\alpha = 5\%$ وهذه العينة من

الإجابات تؤكد عشوائية العينة عند مستوى المعنوية 5% .

ب : اختبار المتتابعات فى حالة العينات الكبيرة :

كان الاختبار السابق فى حالة العينات الصغيرة ، أى حين تكون كل من

n ، n_1 ، n_2 أما فى حالة العينات الكبيرة ، أى حين تزيد قيمة كل من n ،

n_1 عن ٢٠ فإنه يمكن استخدام التقارب الإعتدالى فى اختبار العشوائية إذ أن

توزيع العينات للمتتابعات توزيع معتاد:

$$(8) \quad \text{توقعه } \mu = 1 + \frac{2 \times 1 \times 2}{2 + 1}$$

$$(9) \quad \text{وإنحرافه المعياري } \sigma = \sqrt{\frac{(2 - 1 - 2 \times 1 \times 2)(2 \times 1 \times 2)}{(1 - 2 + 1)(2 + 1)}}$$

وبالتالى فإن المتغير العشوائى : $\frac{\mu - \bar{x}}{\sigma} \rightarrow$ توزيعى م (صفر، ١) (١٠)

وبالتالى يمكن استخدام التوزيع المعتاد المعيارى فى اختبار العشوائية.

مثال (٩):

أعلنت مؤسسة ما عن طلب عاملين من الذكور والإناث للعمل بالعلاقات العامة وكان ترتيب ورود الطلبات المقدمة للمؤسسة بحسب النوع (ذكور : ذ ، إناث : أ) كالتالى :

ذ أ ذ أ ذ ذ ذ ذ ذ أ ذ أ ذ
ذ أ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ
أ أ أ ذ أ ذ أ ذ أ ذ أ ذ أ ذ
ذ أ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ ذ

إختبر هذه السلسلة من حيث العشوائية عند $\alpha = 1\%$.

الحل:

ن_١ = ٣٠ ذكور ، ن_٢ = ٢٠ إناث

الفرض الأسمى : أن الطلبات عشوائية فى تسلسل ورودها عند $\alpha = 1\%$
ويطبق اختبار المتتابعات للعينات الكبيرة ويرفض الفرض الأسمى إذا كانت :

$$\text{س} = \frac{\mu - \bar{x}}{\sigma} = - > 2.58 \text{ أو } 2.58 < \mu - \bar{x}$$

ومن بيانات العينة

$$35 = \bar{x}$$

$$s^2 = \frac{1}{n} \left(\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right) = \frac{1}{20} \left(1200 + \frac{20 \times 30 \times 2}{20 + 30} - 1 + \frac{2 \times 12}{20 + 12} - 1 + \frac{2 \times 12}{20 + 12} \right) = 11.265$$

$$s = \sqrt{11.265} = 3.356$$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{35 - 30}{\frac{3.356}{\sqrt{20}}} = 2.58 < 2.98$$

ويرفض الفرض بعشوائية الطالبات المقدمة بحسب النوع وذلك بحسب تسلسل

ورودها عند $\alpha = 1\%$

ملحوظة:

ويلاحظ انه بسبب طبيعة القياسات وانها ترتيبية (بحسب نتائجها) فإنه لا يوجد اختبار معلمى أو بارامترى يصلح للاستخدام كبديل لاختبار المتتابعات لذلك فليس هناك مبرر عن قوة اختبار المتتابعات.

(3) اختبار الإشارة Sign Test:

١- الاختبار فى حالة عينة واحدة:

يستخدم اختبار الإشارة لعينة واحدة لاختبار الفروض الإحصائية المتعلقة بمقياس للنزعة المركزية (Central Tendency) فى المجتمع سواء كان هذا المقياس هو الوسط الحسابى μ أو الوسيط ولذلك لم نقل بأنه اختبار لبارامترى بشأن الوسط الحسابى. ويفترض هنا أن العينة سحب عشوائية من مجتمع تتبع فيه الظاهرة موضوع الاختبار توزيعاً مستمراً ولكننا لا نفترض فرضاً ما بشأن

شكل التوزيع ولذلك فهو اختبار بديل لاختبارات الخاص بالوسط الحسابي للمجتمع μ الذي يفترض أن يكون التوزيع في مجتمع العينة معطاداً. وحين يكون التوزيع في مجتمع الدراسة مستمراً ومتماثلاً أو قريباً من التماثل بحيث يكون احتمال أن تأخذ أحد مفردات العينة العشوائية قيمة ما تقل عن الوسط الحسابي μ مساو لاحتمال أن تأخذ قيمة أكبر من μ وهذا الاحتمال $\frac{1}{2}$. وفي هذه الحالة فإن مقياس النزعة المركزية موضوع الاختبار باستخدام اختبار الإشارة هو الوسط الحسابي في المجتمع μ أو الوسيط في المجتمع أيهما يصلح لهذا الفرض. أما إذا كان التوزيع ملتوياً فإن مقياس النزعة المركزية موضوع الاختبار سيكون الوسيط في المجتمع اعتماداً على أن تعريف الوسيط بأنه القيمة التي تتوسط التوزيع وأن يكون عدد المفردات الأكبر منها في القيمة مساو لعدد المفردات الأصغر في القيمة.

هذا بالإضافة إلى عدم تأثير الوسيط بالقيم القليلة المتطرفة كالوسط الحسابي. وأما إذا كان التوزيع متماثلاً أو قريباً من التماثل فإن الوسط الحسابي والوسيط يكونا متساويين في القيمة أو يقتربان من التساوي ولذلك فيصلح أيهما للإستخدام مع اختبار الإشارة.

خطوات الاختبار :

- (١) الفرض العنمی هو $\mu = \mu_0$ أو أن الوسيط $(\mu_r) =$ قيمة معينة مقابل فرض بديل مناسب.
- (٢) يحدد مستوى المعنوية المناسبة α .
- (٣) ثم تستبدل قيم العينة التي نقل عن μ_0 بالإشارة (-) والقيم التي تزيد عن μ_0 (أو الوسيط μ_r) بالإشارة (+). وبذلك يكون الاختبار هو اختبار لمتغير له توزيع ذو الحدين حيث $L = \frac{1}{2}$. وتستبعد المفردات التي تتساوى في القيمة مع μ_0 .

أ: اختبار العينة في حالة العينات الصغيرة:

إذا كان حجم العينة $n \geq 20$ اعتبرت العينة صغيرة وبالتالي فيمكن تحديد احتمال مشاهدة الإشارات (+) أو (-) بعدد مساو للعدد المشاهد أو أكبر منه أو أقل منه ويقبل للفرض العدمي أو يرفض حسب الأحوال ويمكن استخدام اختبار الإشارة اعتماداً على توزيع ذو الحدين كما يوضحه المثال التالي:

مثال (۱۰) :

البيانات التالية تمثل متوسط عدد أيام الأجازات التي حصل عليها عمال المبيعات في عينة عشوائية مكونة من ١٥ فرعاً من أفرع أحد المؤسسات التجارية وذلك خلال أحد السنوات:

اختبر الفرض $\mu = 32$ يوماً عند $\alpha = 0.01$.

سوف يحل المثال باستخدام اختبار الإشارة لعينة واحدة ثم بفرض أن كافة شروط استخدام اختبار قائمة فسوف نعيد المثال باستخدام اختبار ت:

الحل:

(١) اختبار الإشارة:

بإستبدال القيم بالإشارات (+) لتلك التي تزيد فى قيمتها عن ١. ، ٣٢ ، (-) لتلك التي تقل عن هذه القيمة وإستبعاد المفردة الثالثة عشرة وتساوى ١. = ٣٢ تكون الإشارات:

++++-++++ وعدها ١٤ إشارة.

(-) ۱۲ ، (+) ۱۲

وإذا كان الفرض العدمي $\mu = 32$

والبديل $\mu < 32$

ويتوافر في حالتنا هذه كافة الشروط الخاصة بالتوزيع ذو الحدين: وهي أن كل محاولة ينتج عنها أحد ناتجين إما نجاح (عدد أيام الأجازات < ٣٢ يوماً) أو فشل وهو العكس ، وكل باحتمال ثابت (ل = ٠,٥) وعدد المحاولات ن ثابت وهو ١٥ فى مثالنا هذا ، المحاولات المستقلة لذلك فإنه يمكن حساب احتمال الحصول على ١٢ (+) أو ١٣ (+) أو ١٤ (+) وهو ما يعادل الحصول على ٢ (-) أو أقل. وباستخدام ذو الحدين:

$$ح (س \geq ٢) = \sum_{س=٢}^{١٥} \binom{١٥}{س} \left(\frac{١}{٢}\right)^س \left(\frac{١}{٢}\right)^{١٥-س}$$

$$= \binom{١٥}{٢} \left(\frac{١}{٢}\right)^٢ \left(\frac{١}{٢}\right)^{١٣} + \binom{١٥}{٣} \left(\frac{١}{٢}\right)^٣ \left(\frac{١}{٢}\right)^{١٢} + \dots + \binom{١٥}{١٥} \left(\frac{١}{٢}\right)^{١٥} \left(\frac{١}{٢}\right)^٠$$

ويمكن الحصول على قيمة هذا الاحتمال من ملحق رقم (٩) الذى يعطى احتمال الحصول على س أو أقل إشارة (الإشارة الأقل عدداً وهي (-) فى مثالنا) حيث ل = م = $\frac{١}{٢}$ وإلا استخدمت الصيغة.

$$مج (س) = \sum_{ل=٠}^{ن} \binom{ن}{ل} \left(\frac{١}{٢}\right)^ل \left(\frac{١}{٢}\right)^{ن-ل} \quad (١١)$$

للحصول على الاحتمال المطلوب متى كانت ل ≠ م ≠ $\frac{١}{٢}$

وفى مثالنا هذا:

$$ح (س = ١٠, ١١, ١٢ | ن = ١٤) = ٠,٠٠٦ + ٠,٠٠١$$

$$ح (+) = ١٢ إشارة أو أكثر$$

$$= ٠,٠٠٧ > ٠,٠١ \text{ لذلك يرفض الفرض العنمى.}$$

ملحوظة:

إذا كان الفرض العنمى لم = ٣٢ يوماً.

والفرض البديل لم ≠ ٣٢ ، > ٣٢ أو < ٣٢ أى اختبار طرفين.

فإنه يمكن تكرار نفس خطوات الاختبار السابق مع مضاعفة احتمال

الحصول على:

$$(س) = ٢٠١٠٠ | ن = ١٤ = ٠,٠٠٧ \times ٢ = ٠,٠١٤$$

اختبار ت:

$$\bar{م} = \frac{٥٣٩,٨}{١٥} = ٣٥,٩٨٧ \text{ يوماً}$$

$$ع = \frac{٥٠٤,١١}{\sqrt{١٤}} = ٦,٠٠١$$

الفرض العدمي $\mu = ٣٢$ يوماً الفرض البديل $\mu < ٣٢$

$\alpha = ٠,٠١$ وعند $\alpha = ١\%$ يرفض الفرض العدمي إذا كانت $t^* \leq ٢,٦٢٤$

ومن بيانات العينة:

$$t^* = \frac{\bar{م} - \mu}{\frac{ع}{\sqrt{ن}}} = \frac{٣٥,٩٨٧ - ٣٢}{\frac{٦,٠٠١}{\sqrt{١٥}}} = ٢,٥٧٤$$

لذلك يقبل الفرض العدمي عند $\alpha = ١\%$.

واضح من هذا أن القرار قد اختلف في حالة استخدام الإشارة في اختبار ت حيث كان إجراء اختبار الإشارة (أو ما يشار إليه أحياناً باسم اختبار ذو الحدين The Binomial Test) على أساس استبدال القيم بإشارات موجبة وسالبة الذي حول التوزيع إلى توزيع ذو الحدين.

ومع ذلك فإنه ما لم تكن القراءات وصفية أو نوعية مما لا يصلح معها استخدام اختبار ت فإن اختبار الإشارة يكون أقل قوة إذ تصل قوته المكافئة إلى ٩٥٪ وتتناقص حتى تصل إلى ٦٣٪. وإذا كانت القراءات يصلح معها تطبيق اختبار ت فإنه يمكن اختبار مطابقة توزيعها للتوزيع المعتاد - باستخدام اختبار كاس كما أوضحنا في بداية هذا الفصل - وإذا تبين أن توزيعها يتبع التوزيع المعتاد أو لا يختلف كثيراً عنه فإنه يمكن استخدام اختبار ت وإلا استخدم اختبار الإشارة.

٢: اختبار الإشارة في حالة العينات الكبيرة:

إذا كانت $n < 20$ اعتبرت العينة كبيرة وبدلاً من استخدام التوزيع ذو الحدين لإجراء الاختبار فإنه يمكن الاعتماد على التقارب الإعتدالي للتوزيع ذو الحدين متى توافر الشرط اللازم وهو أن n م أو n ل كلاهما أكبر من 5 في هذه الحالة فإن أداة الاختبار هي:

$$Y \sim \frac{S - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{S - (n/2 \pm 0.5)}{\sqrt{n/4}} \quad (12)$$

وقد استخدمت العلامة \sim للدلالة على "تؤول إلى". أما n ل ، n م فهي توقع وتباين S وهو متغير عشوائي له توزيع ذو الحدين - وليس نسبة - أما ل ، م. المعرفتين يفرض العدم.

ويمكن استخدام أي من الصيغتين السابقتين الأولى بدون تقريب أو تصحيح للإستمرارية والثانية مصححة حيث $S + 0.5$ إذا كانت $S > n/2$ ، $S - 0.5$ إذا كانت $S < n/2$.

مثال (١١) (٥):

البيانات التالية توضح كمية العادم اليومي من أكسيد الكبريتيك الذي يخرج من أحد الوحدات الصناعية الكبيرة (بالطن):

٩	٢٧	٢٥	٢٢	١٨	١٩	٢٩	٢٠	١٥	١٧
٢٦	٢٤	٢٣	١٥	١٤	٢٤	٦	١٧	٢٠	٢٤
١٣	٢٠	١٧	٢٤	٢٢	١٦	١٩	٢٨	٢٣	١٩
١٤	٢٤	١٧	٢٠	١٣	٣١	١٨	٢٣	١٠	١٩

استخدم اختبار الإشارة لعينة واحدة لاختبار الفرض بأن $\mu_0 = 23.5$ طن مقابل الفرض بأن $\mu > 23.5$ عند $\alpha = 5\%$.

^{١٥٦} Freund, J. E. & William, F.J. : Elementary Business Statistics : The Modern Approach, Prentice-Hall International Ed., 4th ed., p. 411., 1964.

الحل:

عدد الإشارات + = ١١

عدد الإشارات - = ٢٩

ولأن $N = L = N = M$ $\times 40 = 5 < 20$

لذلك يمكن استخدام اختبار الإشارة (نمو الحدين) مع التقارب الإعتدالي.

$$U = \frac{S - N}{\sqrt{N \cdot L}} = \frac{\frac{1}{2} \times 40 - 11}{\sqrt{\frac{1}{2} \times 40 \times 4}} = -2.85$$

وهي أصغر من -1.645 لذلك يرفض الفرض العنمي.

ب: اختبار الإشارة في حالة عينتين غير مستقلتين:

وينظر هذا الأسلوب اختبار القراءات المزدوجة (Paired Observations) أي حين تسجل القراءتين نفس المفردة. ويكون الهدف في حالتنا هذه هو اختبار مؤثر ما وتكون القراءة الأولى على مفردة العينة قبل استخدام أو إدخال المؤثر الذي يراد قياس أثره ثم تسجل القراءة الثانية على نفس المفردة بعد إدخال المؤثر ويستخدم الفرق بين القراءة الأولى والثانية على مفردات العينة في اختبار معنوية أثر المؤثر موضوع الاختبار أو القياس. وبديهي أن القراءات المزدوجة غير مستقلة لتسجيلها على نفس المفردات. ويمكن استخدام أسلوب بارامترى لاختبار الفرق بين المتوسطين ويعتمد هذا الأسلوب على اختبار ت للفروق بين أزواج القيم كما أوضحنا من قبل ، أما المدخل الآخر فهو اختبار لا بارامترى ويعتمد على اختبار الإشارة وسوف نعالج هذين المدخلين بالمثال التالي:

مثال (١٣) (٥) :

لاختبار مدى فعالية نظام جديد لضبط المرور قد سجل عدد الحوادث التي رصدت عند ثمانية تقاطعات رئيسية اختيرت عشوائياً ، وذلك خلال أربعة أسابيع قبل وبعد استخدام النظام الجديد. وكانت النتائج كالتالي:

قبل	٩	٧	٣	١٦	١٢	١٢	٥	٦
بعد	٥	٣	٤	١١	٧	٥	٥	١

اختبر عند مستوى المعنوية ١٠٪ أن النظام الجديد للمرور لا يختلف عن النظام السابق مقابل الفرض البديل أن أعلى كفاءة. والاختبار باستخدام اختبار الإشارة كالتالي:

الحل:

$$(١) \text{ الفرض } \mu_1 = \mu_2 \text{ مقابل } \mu_1 < \mu_2$$

$$(٢) \alpha = ٠,١$$

(٣) باستخدام اختبار الإشارة ، وتكون الإشارة + إذا كانت القراءة الأولى أكبر من الثانية ، - في حالة العكس.

ويرفض الفرض العدمي إذا كان احتمال الحصول على عدد + ≤ ٦ مرات $٠,١ >$.

*١ Freund, J. E. & William, F.J. & B.M.:
Elementary Business Statistics : The Modern Approach.. 6th ed.,
Prentice-Hall, p. 568. 1994 .

جدول (٩)

عدد حوادث المرور المسجلة عند ٨ تقاطعات رئيسية

بحسب نظام المرور

قبل	بعد	الإشارة	الفرق (ف)	ف
٩	٥	+	٤	١٦
٧	٣	+	٤	١٦
٣	٤	-	١-	١
١٦	١١	+	٥	٢٥
١٢	٧	+	٥	٢٥
١٢	٥	+	٧	٤٩
٥	٥	تستبعد	صفر	صفر
٦	١	+	٥	٢٥
			٢٩	١٥٧

عدد الإشارات + = ٦

ومن جدول ملحق رقم (٩)

ح (س + ٦) = ح (س - ١) = ٠,٠٦٣ > ٠,١

لذلك يرفض الفرض العدمى أى أن النظام الجديد أعلى كفاءة فى ضبط عدد

حوادث المرور عن النظام القديم.

وباستخدام اختبار ت للقرارات المزدوجة:

$$ت^* = \frac{\bar{ف} - \text{صفر}}{\frac{\sqrt{٧,٤١٠٧}}{٨}} = \frac{٣,٦٢٥ - \text{صفر}}{٠,٩٢٦٣} = ٣,٩١٢$$

ولكن ت (٠,١) = ١,٤١٥ > ت^* = ٣,٩١٢

لذلك يرفض الفرض بأن $\mu_1 = \mu_2$ وتتفق نتيجة هذا الاختبار مع نتيجة اختبار الإشارة ، ولكن الاختبار البارامترى أعلى قوة من اختبار الإشارة لذلك يفضل الاختبار الأخير متى توافرت شروط صحة استخدامه.

(4) اختبار مجموع الرتب / اختبار U: مان-ويتني

The Mann-Whitney U Test

سبق أن قدمنا اختبارات بارامترية لاختبار الفرق بين متوسطين $\mu_1 = \mu_2$ ، اعتماداً على وسطين حسابيين \bar{x}_1 ، \bar{x}_2 حسبا من عينتين مستقلتين سواء كان التباين للمجتمعين معلوماً أو غير معلوم وفي الفقرة التالية نقدم اختباراً لا معلمياً (لا بارامترى) يعتمد على الرتب لاختبار الفرق بين المتوسطين لعينتين مستقلتين سحباً كل من مجتمع له توزيع مستمر وليس بالضرورة معتاداً.

وهناك مجموعة الاختبارات على مجموع الرتب المندمجة للعينتين وذلك لاختبار - ما إذا كانت عينتين عشوائيتين مستقلتين قد سحبتا من مجتمع واحد أم لا هذا متى كانت القياسات ترتيبية أو يمكن استخدامها في ترتيب القراءات. بل وبمسيل الباحثون إلى تفضيل استخدام هذا الاختبار بدلاً من اختبارات للفرق بين المتوسطين ($\mu_1 = \mu_2$) متى كانت شروط استخدام اختبارات غير مؤكدة.

خطوات الاختبار:

(1) تدمج مفردات العينتين المستقلتين معاً (مع تمييزهما بحسب العينة التي تنتمي إليها كل) ثم تستبدل المفردات أو المشاهدات بالرتب مرتبة ترتيبياً تصاعدياً أو تنازلياً - مع إعطاء المفردات المكررة المتوسط الحسابي للرتب المناظرة.

(2) وإذا كانت n_1 ، n_2 ترمز إلى عدد مفردات ومجموع رتب العينة الأصغر حجماً ، n_1 ، n_2 إلى عدد مفردات ومجموع رتب العينة الأكبر حجماً فإن أداة الاختبار تتحدد على أساس قيمة n .

أ: اختبار U في حالة العينات الصغيرة:

فإذا كانت $n_2 \geq 20$ استخدم مدخل " العينات الصغيرة " في اختبار الفرض بأن العنيتين تنتميان إلى مجتمع واحد ، أى أن التوزيع واحد للظاهرة في العنيتين. وتستخدم أداة الاختبار.

$$(13) \quad U = n_1 n_2 - \frac{n_1(n_1+1)}{2} + \frac{n_2(n_2+1)}{2} - \frac{(n_1+n_2+1)^2}{2}$$

$$(14) \quad U = n_1 n_2 - \frac{n_2(n_2+1)}{2} + \frac{n_1(n_1+1)}{2} - \frac{(n_1+n_2+1)^2}{2}$$

وتقارن القيمة المحسوبة لـ U بالقيمة المعنوية لـ U_α كما يحددها جدول ملحق رقم (١٠) ويرفض الفرض العدمي إذا كانت $U \geq U_\alpha$ ويقبل فيما عدا ذلك من الأحوال. وعادة تستخدم U الأصغر قيمة في الاختبار. وإذا تبين أن القيمة المحسوبة $U < \frac{n_1 n_2}{2}$ فإنها ستكون الأكبر قيمة ولإيجاد القيمة الأصغر لـ U تستخدم العلاقة التالية:

$$(15) \quad U = n_1 n_2 - U^*$$

فمثلاً إذا حسبت U ، وتبين أنها $\frac{n_1 n_2}{2}$ فبدلاً من حساب U من (١٤) فإنه يمكن إيجاد قيمتها من العلاقة (١٥) أى أن :

$$U = n_1 n_2 - U$$

ب: اختبار U في حالة العينات الكبيرة:

أما إذا كانت $n_2 \leq 20$ فيستخدم المدخل الخاص بالعينات الكبيرة الذى يعتمد على التقارب الاعتدالى لتوزيع المتغير العشوائى U إذ أن توزيع العينات للمتغير العشوائى U يقترب من التوزيع المعتاد الذى متوسطه

$$(16) \quad \mu_U = \frac{n_1 n_2}{2}$$

$$(17) \quad \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{12} = \sigma^2 \quad \text{وتباينه}$$

$$\text{وبالتالى فإن } \frac{U - \mu}{\sigma} \text{ له توزيع يؤول إلى التوزيع}$$

$$(18) \quad \text{المعتاد المعيارى: م (صفر ، ١)}$$

وعليه فتستخدم أداة الاختبار (١٨) فى اختبار الفرض فى هذه الحالة. والأملية التالية توضح كيفية استخدام هذا الاختبار فى حالة العينات الصغيرة والكبيرة.

مثال (١٣):

البيانات التالية تبين الطاقة الحرارية (مليون سعر حرارى / للطن) لعينة عشوائية من الفحم من منجمين مختلفين:

المنجم أ:	٨٣٨٠	٨٢١٠	٧٣٦٠	٧٨٤٠	٧٩١٠
المنجم ب:	٧٥٤٠	٧٧٢٠	٧٧٥٠	٨١٠٠	٧٦٩٠

اختبر الفرض بأن العينتين من مجموعتين متماثلتين عند $\alpha = ٥\%$.

الحل:

الفرض العدمى: أن العينتين من منجمين متماثلتين تماماً أما الفرض البديل فهو أن المنجمين يختلفان فى متوسط الطاقة الحرارية للفحم المستخرج.

$$\alpha = ٥\%$$

وبترتيب المفردات فى العينتين معاً ترتيباً تصاعدياً نجد أن:

٨٣٨٠	٨٣٦٠	٨٢١٠	٨١٠٠	٧٩١٠	٧٨٤٠	٧٧٥٠	٧٧٦٠	٧٦٩٠	٧٥٤٠	ج
ب	ب	ب	ب	أ	أ	ب	ب	ب	ب	منجم
١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	ترتيب

$$٣٨ = ر \quad ١٧ = ب$$

$$U^* = \frac{n_1}{n} + \frac{n_2}{n} - \frac{(n_1 + n_2)^2}{n} = 0.5 - \frac{0.5^2}{1} = 0.25$$

$$12.5 = \frac{25}{2} - \frac{25^2}{2} < 23 = 17 - \frac{6 \times 5}{2} + 5 \times 5 =$$

وبالتالى فإن U^* هي الأكبر قيمة ولحساب قيمة U^* (أى الأصغر قيمة) باستخدام العلاقة (١٥).

$$U^* = 23 - 25 = 2$$

وهى $U = 2\%$ طرفين كما هو مبين بجدول ملحق رقم (١٠) لذلك يرفض الفرض بعدم وجود اختلاف فى التوزيع فى المجتمعين. وجدير بالذكر أنه باستخدام (١٣) فإن:

$$U^* = \frac{n_1}{n} + \frac{n_2}{n} - \frac{(n_1 + n_2)^2}{n} = 0.5 - \frac{0.5^2}{1} = 0.25$$

ولو طبقنا اختبار ت على بيانات نفس المثال حيث:

الفرض العدمى $\mu_1 = \mu_2$ والفرض البديل $\mu_1 \neq \mu_2$. وعند $\alpha = 5\%$ فإن:

$$t^* = 2.61 > 2.305 = t_{0.05, 1}$$

لذلك يقبل الفرض العدمى بعدم اختلاف الطاقة الحرارية المتوسطة للفحم المستخرج من المنجمين. وفى الحقيقة فإن التباين فى العينة الأولى حوالى مرة ونصف التباين فى العينة الثانية مما يعيب استخدام اختبار ت.

مثال (١٤):

سحبت عينتين عشوائيتين الأولى من عمال قسم المبيعات فى صناعة الكيماويات والأخرى عمال قسم بيع أدوات منزلية وسجلت الأجور اليومية لأفراد كل من العينتين فكانت النتائج كالتالى:

جدول (١٠)

الأجور اليومية لعينتين كل من عمال المبيعات في قسمين مختلفين

العينة الأولى:

٦,٩٠٠	٨,٧٠٠	٧,٣٠٠	١٢,٥٠٠	١٠,٤٠٠	٧,٦٠٠
			٧,٨٠٠	١٤,٩٠٠	٩,٧٠٠

العينة الثانية:

٨,٥٠٠	٩,٩٠٠	٨,٣٠٠	١٠,١٠٠	٩,٢٠٠	٨,٨٠٠
	٩,١٠٠	٩,٤٠٠	٩,٠٠٠	٧,١٠٠	١١,١٠٠

اختبر الفرض بعدم وجود فرق بين الأجر اليومي لعمال القسمين.

الحل:

تدمج بيانات العينتين معاً ثم ترتب المفردات ترتيباً تصاعدياً (أو تنازلياً) مع إعطاء متوسط الرتب للمفردات المكررة فمثلاً:

جدول (١١)

الترتيب التصاعدي لبيانات جدول (١٠) للعينتين

٨,٣٠٠	٧,٨٠٠	٧,٦٠٠	٧,٣٠٠	٧,١٠٠	٦,٩٠٠
٢	١	١	١	٢	١
٩,٢٠٠	٩,١٠٠	٩,٠٠٠	٨,٨٠٠	٨,٧٠٠	٨,٥٠٠
٢	٢	٢	٢	١	٢
١١,١٠٠	١٠,٤٠٠	١٠,١٠٠	٩,٩٠٠	٩,٧٠٠	٩,٤٠٠
			٢	١	٢
٢	١			١٤,٩٠٠	١٢,٥٠٠
				١	١

(استخدم الرقم ١ للإشارة إلى العينة الأولى والرقم ٢ إلى العينة الثانية)

وعلى ذلك فإن مفردات كل من العينتين تحتل الرتب التالية

الأولى	١	٣	٤	٥	٨	١٤	١٧	١٩	٢٠	
الثانية	٢	٦	٧	٩	١٠	١١	١٢	١٣	١٥	١٨

وإذا صح فرض عدم وجود فرق بين العينتين فإنه يتوقع أن يكون مجموع رتب مفردات كل من العينتين متقارباً أو متساوياً.

ولأن $n_1 = 11$ و $n_2 = 20$ لذلك يستخدم أسلوب اختبار العينة الصغيرة.

$$U_1 = 11 \times 9 - \frac{10 \times 9}{2} - 91 = 53 - \frac{90}{2} = 53 - 45 = 8$$

وهي الأكبر قيمة ولإيجاد U الأصغر قيمة وهي U_2

$$U_2 = 20 \times 11 - \frac{11 \times 20}{2} - 119 = 220 - 110 - 119 = 91$$

$$U = \min(U_1, U_2) = \min(8, 91) = 8$$

وبمقارنة U مع U_{α} أي 23 أي أن $U < U_{\alpha}$

لذلك يقبل الفرض بعدم وجود اختلاف معنوي بين الأجر اليومية للعمال في القسمين بمعنى أنهما ينتميان إلى مجتمع واحد أو مجتمعين متساويين.

مثال (١٥):

الجدول التالي (١٢) يبين ترتيب مفردات عينتين عشوائيتين من الذكور والإناث في اختبار للقدرة الميكانيكية.

جدول (١٢)

الذكور	الإناث
٢٦	٧
١٠	٢٠
٣٠	٢٧
٣	٢٨
٣٢	٤
٣٨	٣٣
٢٤	١
٤٠	٩
٢٥	١٣
١٢	٢١
٦	٣٩
١٤	٢
١٧	٥
٢٢	٣٦
٣٤	٨
	١١
	١٦
	١٩
	٢٩
	٣١
	١٨
	٣٥

هل تبين هذه الرتب اختلافات معنوية فى القدرات الميكانيكية بين الذكور والإناث عند مستوى المعنوية ٥% ؟

الحل:

ن_١ : للذكور = ١٥ ن_٢ : للإناث = ٢٥

ن < ٢٠ يستخدم التقارب الإعتدالى لـ U فى الاختبار.

$$U_{\mu} = \frac{2 \times 15}{2} = \frac{30}{2} = 15$$

$$U_{\sigma} = \frac{(2 \times 15) (1 + 2 \times 15)}{12} = \frac{450}{12} = 37.5$$

$$U_{\sigma} = \frac{(41) \times 15}{12} = \frac{615}{12} = 51.25$$

$$U_1^* = \frac{(1+15) \times 15}{2} + 2 \times 15 = \frac{225}{2} + 30 = 112.5 + 30 = 142.5$$

$$333 - \frac{16 \times 15}{2} + 25 \times 15 = 333 - 120 + 375 = 588$$

$$162 = 333 - 120 + 375 = 588$$

$$\therefore U^* = \frac{142.5 - 162}{37.5} = -0.512 < -1.96 < 1.96$$

لذلك يقبل الفرض بعدم وجود اختلاف معنوى بين القدرات الميكانيكية للذكور والإناث عند $\alpha = 5\%$.

وجدير بالذكر أن استخدام U بدلاً من U_١ فى الاختبار سوف لا يؤثر إلا على إشارة U* فتتغير من - إلى + دون أن يؤثر ذلك على القرار بشأن الفرض موضوع الاختبار ويتأكد ذلك من الآتى:

$$U_2^* = \frac{(2 \times 15) (1 + 2 \times 15)}{12} + 2 \times 15 = \frac{450}{12} + 30 = 37.5 + 30 = 67.5$$

$$213 = 487 - 320 + 370 =$$

$$213 = 162 - 370 = {}^*U - 1 \text{ ن} =$$

$$\therefore \text{ى}^* = \frac{187,5 - 213}{35,795} = 0,712 < 1,96 < 1,96$$

ويقبل الفرض العدمى عند $\alpha = 5\%$.

وهذا يؤكد أيضاً الإتجاه إلى *U الأكل قيمة فى الاختبار.

ج: ملاحظات على اختبار U :

- (١) يفترض لجواز استخدام هذا الاختبار - كبديل للاختبار البارامترى ت للفرق بين متوسطين - أن يكون توزيع المتغير فى المجتمعين مستمراً وليس بالضرورى أن يكون معتاداً. وأن تكون العينتين مستقلتين.
- (٢) ويفضل استخدام هذا الاختبار بدلاً من اختبار t إذا كان واضحاً من البداية أن العينتين مختلفتى التباين بشكل واضح مما لا يصلح معه استخدام اختبار t الذى يفترض أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ وبالتالى يمكن تقدير σ^2 بالتباين التجميعى \bar{S}^2 من بيانات العينتين. كما يفضل أيضاً إذا كان الحصول على قيم دقيقة للملاحظات غير ميسور وإن كان من الممكن ترتيبها.

- (٣) إن مجموع رتب العينتين أى $r_1 + r_2 =$ مجموع الـ $n_1 + n_2$ الأولى من سلسلة الأعداد الطبيعية وتساوى $\frac{(n_1 + 1)(n_2 + 1)}{2}$ ومن هذه العلاقة فإنه يمكن معرفة r_1 من حساب r_2 والعكس صحيح ولذلك يمكن إجراء الاختبار على أساس المقياس الإحصائى U لأى من العينتين خاصة إذا كانت $n_1 = n_2$ وإلا فإنه يفضل -

تسهيلاً للعمليات الحسابية - حساب U لأصغر العينتين حجماً.

- (٤) إن مجموع $U_1 + U_2 = n_1 n_2$ وفى الحقيقة فإن توزيع العينات لـ U متماثل حول $n_1 n_2 / 2$.

(٥) تستبدل الرتب المكررة بالوسط الحسابي لها ولا يؤثر ذلك على قيمة المقياس U^* ما لم تكن الرتب المكررة تشمل مجموعتي البيانات المقارنة. وفي الحالة الأخيرة فإنه قد ينصح باستخدام تصحيح معين يؤدي إلى زيادة محدودة في قيمة U^* يمكن إهمالها.

(٥) اختبارات ولوكسن التي تعتمد على الرتب:

١- اختبار ولوكسن لمجموع الرتب: اختبار W :

The Wilcoxon Rank-Sum Test

وهو اختبار مان-ويتني لمجموع الرتب يستخدم مجموع الرتب لاختبار ما إذا كان مجتمعين مستقلين مستمرين غير مختلفين. أم أن أحدهما يختلف عن الآخر وذلك اعتماداً على بيانات عينتين مستقلتين تسحب كل من أحد المجتمعين والثانية من المجتمع الآخر وهو بالتالي يعتبر البديل اللابارامترى لاختبار ت للفرق بين متوسطين. وينسب هذا الاختبار إلى Frank Wilcoxon (١٨٩٢-١٩٦٥) وسوف نطلق عليه اختبار W كما يعرف به في المراجع الإحصائية. وهناك جداول خاصة تستخدم في حالة العينات الصغيرة ولكننا بدافع السهولة سوف نقصر على تقديم هذا الاختبار في حالة العينات الكبيرة (ن، ن٢) كلاهما < ١٠). وخطوات الاختبار هي ذاتها لاختبار U خطوات الاختبار:

(١) تدمج مفردات العينتين المستقلتين (أ، ب) مرتبة ترتيباً تصاعدياً مع تمييز رتب كل من العينتين مع إعطاء المفردات ذات نفس الترتيب الوسط الحسابي للرتب المكررة.

(٢) تجمع رتب كل من العينتين د، د٢.

(٣) واعتماداً على مجموع رتب أي من العينتين، وإذا كان حجم كل من العينتين لا يقل عن ١٠ مفردات فإنه يمكن استخدام اختبار W الآتي:

W : مجموع رتب العينة أ أي د.

وإذا صح فرض العدم بأن المجتمعين متماثلين وغير مختلفين فإن W أو D لها توزيع عينات توقعه:

$$(19) \quad \frac{n_1(n_1 + n_2 + 1)}{2} = {}_{1D}\mu = {}_W\mu$$

وتباينه:

$$(20) \quad \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = {}_{1D}\sigma^2 = {}_W\sigma^2$$

وبالتالي فإن أداة الاختبار هي:

$$(21) \quad U = \frac{{}_{1D}\mu - {}_{1D}\mu}{\frac{{}_{1D}\sigma^2}{W}} = \frac{{}_W\mu - W}{\frac{{}_W\sigma^2}{W}}$$

وهذا المتغير العشوائي له توزيع U : م (صفر ، ١) وبالتالي يمكن إجراء الاختبار في الاتجاه واحد أو اتجاهين وفق ما يمليه الفرض البديل وذلك بالمقارنة بقيم U المعنوية كالمعتاد.

مثال (١٦):

استخدم بيانات مثال (١٥-٥) لاختبار الفرض فإن مجتمع المنكور ومجتمع الإنث كلاًهما متماثلان عند $\alpha = 5\%$.

الحل:

$$D = W = 333$$

$$307,5 = \frac{(1+25+15) \cdot 15}{2} = {}_{1D}\mu = {}_W\mu$$

$$1281,25 = \frac{(1+25+15) \cdot 25 \cdot 15}{12} = {}_{1D}\sigma^2 = {}_W\sigma^2$$

$$30,7946 = {}_{1D}\sigma = {}_W\sigma$$

الفرض : متوسط القدرات الميكانيكية في المجتمعين متساويان

(أي أن $\mu_1 = \mu_2$).

الفرض البديل: $\mu \neq \mu_0$ أى أنه إما $\mu > \mu_0$ أو $\mu < \mu_0$.
 $n = 10$, $n_p = 25$ وبالتالي يمكن استخدام اختبار العينات الكبيرة Y
 حيث:

$$Y = \frac{\bar{d} - \mu_0}{\sigma_d}$$

وإذا صح فرض العدم فإن المتغير العشوائى له توزيع Y . ويرفض فرض العدم إذا كانت $Y > 1.96$ أو $Y < -1.96$ عند $\alpha = 5\%$.

الحل:

ومن البيانات:
 $Y^* = \frac{307.5 - 333}{35.7946} = -0.712$ أو $-1.86 < -0.712 < 1.96$
 . يقبل الفرض بتمثل المجتمعين من حيث القدرات الميكانيكية المتوسطة.
 هذا وإذا أخذت D فى الاعتبار للاختبار بدلاً من D فإن:

$$D = W = 333$$

$$\mu_{Dp} = \frac{n_p(n+1)}{2} = 512.5$$

$$\sigma_D = \sigma_{Dp} = 35.7946$$

$$Y^* = \frac{512.5 - 487}{35.7946} = 0.712$$

ويقبل الفرض العدمى أيضاً، وهو ما كان متوقفاً إذ أن الاختبار فى إتجاهين وبالتالي فالفرض البديل واحد فى الحالتين.
 كما يلاحظ أننا وصلنا إلى ذات القيمة لـ Y^* وبالتالي نفس القرار باستخدام اختبار U :

ب: اختبار ولكوكسن لرتب الفروق بالإشارات: اختبار ر أو (T):

The Wilcoxon Signed-Rank Test

ويعتمد هذا الاختبار على إشارات فروق الرتب وهو في هذا يشبه اختبار الإشارة للقراءات المزدوجة ولكنه لاستخدامه لفر أكبر من المعلومات عن اختبار الإشارة حيث يدخل في الاعتبار ليس إتياء الفروق فقط ولكن حجم الفروق بين القراءات المزدوجة أيضاً ، لذلك فإنه أكفاً من اختبار الإشارة وبالتالي يفضل استخدامه بدلاً من اختبار الإشارة للفروق بين القراءات المزدوجة وسيتضح ذلك من المثال التالي:

خطوات الاختبار:

- (١) تحسب الفروق بين كل زوج من أزواج القراءات المتناظرة فر (ر = ١ ، ٢ ، ... ، ن) مع استبعاد الفروق التي تساوى الصفر ويخفض حجم العينة مقابل كل زوج مستبعد من المشاهدات.
- (٢) ترتب الفروق المطلقة |فر| تصاعدياً ثم يرصد لكل رتبة الإشارة الخاصة بالفروق الذي تناظره هذه الرتبة.
- (٣) تجمع الرتب أخذاً في الاعتبار الإشارات وإذا رمزنا إلى هذا المجموع بالرمز ر (وقد استخدم ولكوكسن الرمز T للدلالة على هذا المجموع).
- (٤) وإذا كان حجم العينة $n \leq 10$ ومتى صح فرض العدم أى تساوى توزيع الظاهرة فى المجتمعين فإنه يتوقع أن يكون مجموع الفروق بإشارات موجبة - مجموع الفروق بإشارات سالبة ، وبالتالي فإن هذا المتغير له توزيع عينات توقعه:

$$\begin{aligned} \mu_r &= \text{صفر} \\ \sigma_r^2 &= \frac{n(n+1)(1+n^2)}{6} \end{aligned} \quad \begin{aligned} (22) \\ (23) \end{aligned}$$

وبالتالى فإن المتغير العشوائى:

$$Y = \frac{N(1+N)(1+2N)}{6} - \text{صفر}$$

(٢٤)

له توزيع Y : م (صفر ، ١)

واعتماداً على ذلك يمكن إجراء الاختبار بتساوى المجتمعين المتناظرين.

مثال (١٧):

بعد أن أدخلت وزارة البترول تعديلاً على البنزين وأصبح البنزين فى بعض المحافظات خالياً من الرصاص ولاختبار ما إذا كان إستهلاك البنزين / اللتر لم يختلف بعد إدخال هذا التعديل عنه قبل ذلك فقد أجريت تجربة النوعين على عينة من ١٥ سيارة مختلفة الطراز وعلى نفس الطريق باستخدام البنزين الخالى من الرصاص أ ومرة أخرى بعد استخدام البنزين العادى ب ، وقد تم تخصيص نوع البنزين عشوائياً على السائقين ، ومن خصص له النوع أ فى الدورة الأولى خصص له النوع ب فى الدورة الثانية وبالعكس. وكانت نتيجة التجربة كالاتى:

طراز السيارة	كيلومتر / اللتر		إشارة ف = أ-ب	ف	ف	رتبة فر	الترتيب بالإشارات
	بنزين خالى من الرصاص (أ)	بنزين عادى (ب)					
١	٢٢,١	٢٣,٧	-	١,٦-	١,٦	٩	٩-
٢	١٥,٧	١٦,١	-	٠,٤-	٠,٤	٣	٣-
٣	١٨,٢	١٩,٠	-	٠,٨-	٠,٨	٥	٥-
٤	١٩,٠	١٩,٠	صفر	صفر	صفر	—	—
٥	٢٥,٧	٢٤,٢	+	١,٥+	١,٥	٧,٥	٧,٥+
٦	١٩,٢	٢٢,٠	-	٢,٨-	٢,٨	١٢,٥	١٢,٥-
٧	١١,٩	١٢,١	-	٠,٢-	٠,٢	١	١-

١٠,٥-	١٠,٥	٢,٢	٢,٢-	-	٣٠,٢	٢٨,٠	٨
١٢,٥-	١٢,٥	٢,٨	٢,٨-	-	٣٧,٨	٣٥,٠	٩
١٤-	١٤	٣,٠	٣,٠-	-	٣٠,٠	٢٧,٠	١٠
٤+	٤	٠,٧	٠,٧+	+	١٩,٠	١٩,٧	١١
٦+	٦	١,٢	١,٢+	+	١٥,١	١٦,٣	١٢
١٠,٥-	١٠,٥	٢,٢	٢,٢-	-	٢٣,٤	٢١,٢	١٣
٧,٥+	٧,٥	١,٥	١,٥+	+	٢٠,٧	٢٢,٢	١٤
٢-	٢	٣,٠	٠,٣-	-	٢٨,١	٢٧,٨	١٥

مجر - ٥٥

الفرض العدمي : متوسط المسافة المطوعة (كيلومتر / باللتر) باستخدام نوعين البنزين أ ، ب واحد.

الفرض البديل : إن متوسط المسافة المقطوعة (كيلومتر / لتر) في حالة استخدام البنزين النوع أ أعلى منه باستخدام النوع ب .

مستوى المعنوية $\alpha = ٥\%$.

باستخدام اختبار الإشارة س = ٤ .

ح (س : + : ≥ ٤ | ن = ١٤)

$$= ٠,٠٠١ + ٠,٠٠٦ + ٠,٠٢٩ + ٠,٠٠٩ = ٠,١٢٦ < ٠,٠٥$$

لذلك يقبل فرض العدم باستخدام اختبار الإشارة.

وباستخدام اختبار ولكوكسن للرتب بالإشارات:

$$\sigma_r = \sqrt{\frac{١٤(١+١٤)(١+٢٨)}{٦}} = ٣١,٨٦$$

$$\therefore \text{حي}^* = \frac{٥٥-}{٣١,٨٦} = ١,٧٢٦ > ١,٦٤٥$$

لذلك يقبل فرض العدم باستخدام اختبار ولكوكسن للرتب بالإشارات ويمكن أن يجرى الاختبارين كل في اتجاهين أى أنه باستخدام اختبار الإشارة
 ح (س : + | ١.٤ ≥ ن = ٤) = ٠.١٢٦ × ٢ = ٠.٢٥٢ < ٠.٢٥
 وبالتالي يقبل فرض العدم. وكذلك بالنسبة لاختبار ولكوكسن للرتب بالإشارات:

ي* = ١.٧٢٦ - > ١.٩٦ > ١.٩٦ أى يقبل فرض العدم أيضاً.

وهناك مدخل آخر لاختبار ولكوكسن للرتب بالإشارات - كما أوضحنا في الفقرة السابقة ويعتمد في ذلك على توزيع مجموع الرتب الموجبة T_+ أو T_- أو مجموع الرتب السالبة T_- أو T_+ ، وإذا صح فرض العدم فإنه يتوقع أن تكون $T_+ = T_-$. أى أن فرق بين أزواج القيم المتناظرة له نفس الاحتمال لأن يكون فرقاً موجباً أو سالباً، وبالتالي فإن المتغير العشوائى T_+ أو T_- له توزيع عينات توقعه:

$$\mu_{T_+} \text{ أو } \mu_{T_-} = \frac{n(n+1)}{4} \quad (٢٥)$$

وتباينه:

$$\sigma_{T_+}^2 \text{ أو } \sigma_{T_-}^2 = \frac{n(n+1)(n+2)}{24} \quad (٢٦)$$

وبالتالى فإن المتغير العشوائى:

$$(٢٧) \quad \begin{cases} Y = \frac{T_+ - \mu_{T_+}}{\sigma_{T_+}} \\ \text{أو} \\ Z = \frac{T_- - \mu_{T_-}}{\sigma_{T_-}} \end{cases}$$

له توزيع Y : م (صفر ، ١) واعتماداً على ذلك يمكن إجراء الاختبار على أساس T_+ أو T_- .

مثال (١٨):

وتطبيقاً لذلك باستخدام r_+ وباستخدام بيانات المثال السابق (١٧-٥) فإن:

$$r_+ = 25 - \mu_r = 52,5$$

$$r_+ = 15,93 - \sigma_r$$

ويستكمل اختبار الفرض كالاتى:

$$\text{الفرض العدمى : } \mu_r = \mu_p$$

$$\text{الفرض البديل : } \mu_r < \mu_p$$

$$\alpha = 0,05$$

ومن البيانات:

$$t^* = \frac{52,5 - 25}{15,93} = 1,726 < 1,645$$

وبقبل فرض العدم: إذا كان الاختبار فى اتجاه واحد عند $\alpha = 5\%$ كما يقبل

الفرض العدمى إذا كان الاختبار فى اتجاهين حيث $1,96 < 1,726 < 1,96$

عند $\alpha = 5\%$.

وإذا طبق ذات الاختبار على مجموع الرتب السالب r_- أو (T_-) .

$$r_- \text{ أو } (T_-) = 80$$

$$\mu_r \text{ أو } (T_- \mu) = 52,5$$

$$\sigma_r \text{ أو } (T_- \sigma) = 15,93$$

$$t^* = \frac{27,5}{12,93} = 1,726$$

وإذا كان الفرض العدمى $\mu_r = \mu_p$

مقابل الفرض البديل $\mu_r \neq \mu_p$ فى اتجاهين.

ي* = ١,٧٢٦ < ١,٩٦ ، > ١,٩٦ لذلك يقبل الفرض العدمى ، أما إذا كان الفرض البديل هو أن التوزيع الاحتمالى لـ أ قد انحاز إلى الجانب الأيسر من التوزيع الاحتمالى لـ ب أى أن $\mu_A > \mu_B$ وهو اختبار طرف أيسر ولأن $Y^* = ١,٧٢٦ < -١,٦٤٥$ لذلك يقبل فرض العدم.

ملحوظة:

قد لا يتفق القرار بالنسبة للفرض العدمى فى حالة استخدام اختبار الإشارة مع القرار باستخدام اختبار الرتب بالإشارات لولكوكسن لأن الاختبار الأخير يستخدم معلومات أكثر من البيانات وفى حالة كهذه يفضل استخدام الاختبار الأخير الأكثر تمييزاً.

(٦) تحليل التباين فى اتجاه واحد بالرتب / اختبار كروسكال-والس Kruskal-Wallis One-way Analysis of Variance

سبق أن قدمنا اختبار تجانس توزيع ظاهرة فى عدة مجتمعات باستخدام توزيع كاسلوب الاختبار فى حالة تعدد العينات المستقلة. وفى هذه الفقرة سوف نقدم أسلوباً آخر هو تحليل التباين فى اتجاه واحد باستخدام الرتب.

اختبار كروسكال / والس:

وهو كالتحليل فى اتجاه واحد باستخدام تحليل التباين - أى التصميم كامل العشوائية - يهدف إلى اتخاذ قرار فى شأن عدة عينات مستقلة تسحب من عدة مجتمعات وأنها لا تختلف فى توزيع المتغير داخلها أى ما إذا كانت الفروق بين متوسطات العينات تعكس فروقاً بين المجتمعات التى سحبت منها تلك العينات أم لا ؟

ويعتمد هذا الاختبار على الرتب خاصة إذا كانت القياسات ترتيبية أو أن التوزيع المعتاد غير محقق وإذا صح فرض العدم بأن متوسطات (أو الوسيط) للمجتمعات التى سحبت منها العينات متساوية فإن متوسطات الرتب لكل عينة

ستكون متساوية وإذا جمعت الرتب في جميع العينات أو في العينة التجميعية فإن متوسطات الرتب $\frac{1+N}{2}$ وبالتالي فإن القيمة المتوقعة للفروق بين متوسط الرتب لكل عينة والمتوسط العام للرتب التجميعية أى توقع $(\frac{1+N}{2} - \frac{r_o}{n_o})$ = صفر .
 وجددير بالذكر أن شكل التوزيع في المجتمعات غير مطلوب التحقق منه وكل ما يتعين

توافره هو استقلال العينات المسحوبة من المجتمعات المختلفة.

خطوات الاختبار:

- (١) يحدد الفرض العدمى وهو أنه لا يوجد فرق بين توزيع المتغير في المجتمعات موضوع الاختبار أما الفرض البديل فهو أن التوزيعات متباينة.
 - (٢) ويحدد مستوى المعنوية α .
 - (٣) ترتب القياسات في جميع الـ $1, 2, \dots, m$ عينة - والتي سحبت كل من مجتمع الدراسة - مجتمعة وترصد الرتب المناظرة لتلك القياسات مع مراعاة الترتيب التصاعدى ثم تجمع الرتب r_o في كل عينة.
- وتستخدم أداة الاختبار:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^m n_i \left(\frac{1+N}{2} - \frac{r_o}{n_o} \right)^2 \quad (28)$$

وهى صيغة تماثل كاسلر لجودة المطابقة والتي ذكرت في بداية هذا الفصل ويقترب توزيعها من توزيع كاسلر أو الصيغة (٢٩) التالية:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^m \frac{n_i^3}{n_i} - \frac{r_o^2}{(1+N)^3} \quad (29)$$

حيث $1, 2, \dots, m$ عينة n_o : حجم العينة و

ن = مج ن_و ر_و = مجموع الرتب فى العينة و

حيث يتبع المتغير العشوائى هـ توزيع كاي^٢ بدرجات حرية م - ١ متى كان حجم كل العينة ن_و < ٥ . أما إذا كان عدد العينات م = ٣ ، وحجم العينة ن_و ≥ ٥ فى كل عينة فإن تقارب توزيع هـ لتوزيع كاي^٢ يصبح غير محقق ويرجع إلى ملحق (١١) الذى يستخدم المعادلة السابقة (٢٨) لحساب احتمال مشاهدة قيمة لـ هـ ≤ القيمة المحسوبة هـ* لقيم مختلفة لـ ن_و . وعليه فيرفض الفرض العدمى إذا كانت ح (هـ ≥ α) ويقبل فيما عدا ذلك.

وكما تبين من الفقرة السابقة فإن هذا الاختبار يحدد ما إذا كانت مجاميع الرتب متباينة وبالتالي فإن العينات لا يتوقع أن تكون من مجتمع واحد (أو عدة مجتمعات متجانسة).

مثال (١٩):

استخدم أسلوب التحليل فى اتجاه واحد بالرتب (اختبار كروسكال/والس) لتحليل بيانات مثال (٢-٣) بالفصل الثالث.

الحل:

(١)	س _١	٤٧	٥٣	٤٩	٥٠	٤٦	مجم _{رو}
الرتب	١	٢	١٠	٣,٥	٥,٥	١	٢٢
(٢)	س _٢	٥٥	٥٤	٥٨	٦١	٥٢	
الرتب	٢	١٣	١١,٥	١٤	١٥	٩	٦٢,٥
(٣)	س _٣	٥٤	٥٠	٥١	٥١	٤٩	
الرتب	٣	١١,٥	٥,٥	٧,٥	٧,٥	٣,٥	٣٥,٥

الفرض العدمى : م_١ = م_٢ = م_٣ أى أن الآلات الثلاث لا تختلف فى إنتاجيتها.

α = ٠,٠٥ وباستخدام وسيلة الاختبار:

$$\begin{aligned} \text{هـ} - &= \frac{12}{(1+n)} - \frac{6}{1-n} - \frac{2}{(1+n)^3} \\ \text{هـ}^* - &= \frac{12}{16 \times 15} - \left\{ \frac{35,0}{5} + \frac{62,0}{5} + \frac{22}{5} \right\} \\ &= \frac{0,05}{5} - (126,25 + 39,6,25 + 48) \\ &= 8,000 - 48 - 56,000 = \end{aligned}$$

وباستخدام المعادلة (٢٨) فإن:

$$\begin{aligned} \text{هـ}^* - &= \frac{12}{16 \times 15} \times \left\{ \left(8 - \frac{35,0}{5} \right) + \left(8 - \frac{62,0}{5} \right) + \left(8 - \frac{22}{5} \right) \right\} \\ &= (34,02) - 8,000 = \end{aligned}$$

وهي نفس القيمة السابقة.

ومن ملحق (١١)

ح ($8 \leq$) أقل من ٠,٠٠٩ لذلك يرفض الفرض اعدى عند $\alpha = 0,05$.
هذا ويلاحظ أننا لو طبقنا تقارب توزيع هـ لتوزيع كا^٢ فإن:
كا^٢ = هـ = ٨,٠٠٥ < كا^٢ = ٥,٩٩ = ٠,٠٥٠٢ فيرفض الفرض أيضاً.
كما سبق أن رفضنا الفرض باستخدام أسلوب تحليل التباين (راجع الفصل الأول).

ملحوظة:

إذا كانت بعض الرتب مكررة - كما هو الحال في المثال الأخير - فتعطى القيم المكررة الوسط الحسابي للرتب المكررة ثم تصحح هـ^{*} كالآتي:

$$\begin{aligned} \text{هـ}^* \text{ مصححة} - &= \frac{\text{هـ}}{1 - \text{مجموع}} \\ &= \frac{\text{هـ}}{1 - \sum_{i=1}^k \frac{f_i}{n}} \end{aligned}$$

حيث $\sum_{i=1}^k f_i =$ ت^١ - ت^٢ ، ت^٢ - ت^٣ ، ... ، ت^٣ - ت^٤ = عدد المشاهدات المكررة في كل عينة.

وهذا التصحيح يؤدي إلى زيادة قيمة ه* وبالتالي فإنه يفقد قيمته العملية إذا كانت قيمة ه* غير المصححة تؤدي إلى رفض الفرض العدمي. وبالإضافة إلى ذلك فإن كثرة عدد الرتب المكررة يقلل من دقة هذا الاختبار الذي يقوم على فرض إستمرارية التوزيعات.

ملاحظات ختامية على الاختبارات التي تعتمد على عدة عينات مستقلة:

سبق أن قدمنا في الفصل الثالث تحليل التباين كمدخل للتحليل في حالة تعدد العينات. كما قدمنا في هذا الفصل اختبار كاي² لتجانس توزيع ظاهرة في عدة مجتمعات كمدخل آخر للاختبار في حالة تعدد العينات وكان المدخل الأول بارامترى والثاني مدخل لابارامترى ثم عرضنا لمدخل لابارامترى ثان هو تحليل التباين في اتجاه واحد بالرتب لكروسكال واليس للاختبارات في العينات المتعددة ويعتينا أن نعرف على السمات الرئيسية التي تميز كل من هذه المداخل الثلاث:

(١) يلاحظ أننا استخدمنا للاختبارات اللابارامترى لم نفترض الشروط الخاصة بتحليل التباين - راجع الفصل الثالث - من حيث التوزيع الإعتدالي للمتغيرات في المجتمعات واستقلال الأخطاء وتوزيعها الإعتدالي و ... إلخ إكتفاء بفرض توافر شروط أيسر كاستمرارية التوزيع أو إستقلالية العينات المسحوبة من مجتمعات مختلفة.

(٢) يستخدم تحليل التباين في اتجاه واحد لكروسكال - واليس متى كان توزيع المتغير مستمراً والقياسات من النوع الترتيبي على الأقل وكفأته النسبية بالمقارنة بتحليل التباين ٩٥,٥% مما يجعل استخدام اختبار كروسكال/واليس مفضلاً عندما نشكك في عدم تحقق في بعض الفروض الخاصة بتحليل التباين البارامترى مثل عدم تماثل توزيع المتغير في المجتمعات أي تساوي تباينها.

(٣) أما إذا كانت القياسات من النوع الوصفي فلا بديل عن اختبار كاي².

(٧) معامل ارتباط الرتب لسبيرمان واختبار معنويته:

سبق أن عرضنا معامل ارتباط الرتب لسبيرمان والمعرف كالآتي:

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (31)$$

والذى يستخدم لقياس العلاقة الخطية بين متغيرين س ، ص حيث ف = رتبة س-رتبة ص لجميع أزواج القيم المتناظرة للمتغيرين س ، ص. وكما أوضحنا فإنه يمكن أن يستخدم هذا المعامل لقياس العلاقة بين المتغيرين إذا كانت القياسات ترتيبية وفي هذه الحالة لا يصلح معامل الارتباط الخطى البسيط لبيرسون لقياس تلك العلاقة ، كما يمكن استخدامه إذا كانت القياسات يمكن أن تستخدم فى ترتيب قيم س ، ص تصاعدياً أو تنازلياً. كما أوضحنا فى الفصل الرابع من هذا الكتاب كيفية اختبار معنوية معامل الارتباط الخطى البسيط لبيرسون اعتماداً على اختبار معنوية معامل الإنحدار الخطى البسيط $\beta_0 =$ صفر حيث اختبار β_0 يعتمد على أداة الاختبار:

$$\frac{\sum (d_i - \bar{d})^2}{n-1} = \frac{r^2 (n-2)}{1-r^2} \quad (32)$$

متوسط مربعات البواقي

وهذا المتغير العشوائى له توزيع ف بدرجات حرية (١ ، ن - ٢) وجذره التربيعى:

$$\frac{r^2 (n-2)}{1-r^2} \quad (33)$$

له توزيع ت بدرجات ن - ٢ وهذا الأخير هو أداة الاختبار للفرض الإحصائى المتعلق بمعامل الارتباط الخطى البسيط لبيرسون $p =$ صفر مقابل الفرض البديل $p \neq$ صفر وعليه فإن اختبار β_0 هو نفسه اختبار p .

وفى هذه الفقرة سوف نقدم كيفية اختبار معنوية معامل ارتباط الرتب لسبيرمان.

أ: معامل ارتباط الرتب لسبيرمان

مثال (٣٠):

لدراسة العلاقة بين طول مدة التغيب وعدد الوحدات المعيبة فقد أجريت دراسة جمعت من خلالها بيانات عن معدل التغيب الأسبوعي المتوسط (س) فى عينة مكونة من ١٢ أسبوعاً وعدد الوحدات المعيبة المنتجة (ص) خلال الأسابيع الاثنى عشرة فكانت النتائج كما هو فى جدول (١٣) التالى:

جدول (١٣)

معدل التغيب المتوسط وعدد الوحدات المعيبة فى عينتين

الأسبوع	١	٢	٣	٤	٥	٦
معدل التغيب	٧,٣	٦,٤	٦,٢	٥,٥	٦,٥	٤,٧
عدد الوحدات المعيبة المنتجة	٢٢	١٧	٩	٨	١٢	٥
الأسبوع	٧	٨	٩	١٠	١١	١٢
معدل التغيب	٥,٨	٧,٩	٦,٧	٩,٦	١٠,٣	٧,٢
عدد الوحدات المعيبة المنتجة	٧	١٩	١٣	٢٩	٣٣	١٨

أوجد معامل ارتباط الرتب بين س: معدل التغيب الأسبوعي المتوسط ، ص: عدد الوحدات المعيبة المنتجة واختبر معنوية معامل الارتباط.

الحل:

$$\frac{1. \times 6}{143 \times 12} - 1 = \frac{1 \times 6}{(1-1) \times 12} - 1 = 0$$

ولإختبار معنوية معامل الارتباط ρ - صفر

519

وبالرجوع إلى ملحق (١٢) والذي يحدد الحدين الأدنى والأعلى (تقدير p بفترة ثقة) لـ p لعند من أزواج القيم $n \geq 30$ لمستويات معنوية مختلفة نجد أنه عند $\alpha = 1\%$ فإن:

$-0.8182 \leq p \leq 0.8182$ ، ولذلك يرفض الفرض العدمي إذ أن $r = 0.965$ تقع خارج منطقة قبول الفرض بعدم وجود علاقة بين s ، v . وفي حالة العينات الكبيرة أى لعند من أزواج القيم $n < 30$ فإنه يمكن الإعتماد على توزيع العينات لمعامل ارتباط الرتب وهو توزيع يؤول إلى التوزيع المعتاد توقعه (صفر) وتباينه $\frac{1}{n-1}$ لذلك فإنه يمكن استخدام المتغير العشوائى:

$$U = \frac{r - \text{صفر}}{\sqrt{\frac{1}{n-1}}} \sim M(\text{صفر}, 1) \quad (34)$$

لاختبار الفروض الإحصائية المتعلقة بـ p .

مثال (٣١):

فى دراسة اجتماعية عن العلاقة بين مستوى الذكاء للزوج والزوجة فى عينة عشوائية مكونة من ٣٢ من حالات الزواج وجد أن معامل ارتباط الرتب لسبيرمان كان $r = 0.809$ اختبر معنوية هذه العلاقة عند $\alpha = 5\%$.

الحل:

الفرض العدمي $p = \text{صفر}$

الفرض البديل $p \neq \text{صفر}$

$\alpha = 5\%$

ولأن $n = 32$ لذلك يمكن استخدام التقارب الإعتدالى لإختبار الفرض حيث:

$$t = \frac{p - r}{\sqrt{\frac{1}{n-1}}}$$

ويرفض الفرض العدمي إذا كانت $t > 1.96$ أو $t < -1.96$

$$\text{ومن بيانات العينة } t = \frac{0.809 - \text{صفر}}{\sqrt{\frac{0.068 \times 0.809}{31}}} = 4.504$$

لذلك يرفض الفرض العدمي والعلاقة بين مستوى ذكاء الزوج والزوجة علاقة معنوية.

تمارين

- ١- ألقيت ٣ قطع عملة ٢٠٠ مرة ورصد المتغير العشوائي X على عدد مرات ظهور الكتابة في كل رمية مستقلة ($X = 0, 1, 2, 3$) فكانت النتائج كالآتي:

X	٠	١	٢	٣
$n(X)$	٣٤	٦١	٧٢	٧٣

فهل تدل نتائج هذه التجربة على أن قطع العملة غير متحيزة عند مستوى المعنوية ٥٪؟

- ٢- الجدول التالي يوضح توزيع أفراد عينة عشوائية من المتقدمين لشغل بعض الوظائف القيادية بحسب الوقت الذي استغرق في أداء اختبار للصلاحيحة:

الزمن بالدقيقة	عدد الأفراد
٢٤ دقيقة أو أقل	١٥
٢٥ -	٥٠
٣٠ -	٧٥
٣٥ -	٤٠
٤٠ -	١٥
٤٥ دقيقة فأكثر	٥

لختبر عند $\alpha = ٥\%$ أن هذه العينة سحبت من مجتمع معناد $\mu = ٣٢,١$ دقيقة، $\sigma = ٥,٦$

- ٣- تيسيراً على عملاء أحد المصارف لصرف شيكات من حساباتهم عن طريق شبايك خاصة لراكبي السيارات drive-in وحتى يمكن للمصرف أن يحدد عدد الشبايك التي تتفق مع الطلب على صرف شيكات عن هذا

الطريق لتحقيق السيوولة المناسبة فقد رصد المصرف فى أحد فروعده عدد مرات تقدم العملاء بشيكات للصرف فى الدقيةة وكانت نتيجة هذه التجربة فى ٧٥٠ حالة كالآتى:

عدد مرات وصول العميل	١	٢	٣	٤	٥	٦	٧+
بسيارته لصرف شيك فى الدقيةة	٤٦	٢٥٠	٢٦٠	١٠٠	٥٠	٣٤	١٠
عدد العملاء							

والمراد اختيار أن هذا التوزيع يتبع التوزيع الاحتمالى المعروف باسم "توزيع بواسون" عند مستوى المعنوية ٥% .

٤- لدراسة العلاقة بين درجة نجاح العاملين فى عملهم بأحد المؤسسات وما حققه كل منهم فى برنامج تدريبى فقد سجلت البيانات التالية من عينة عشوائية من ٤٠٠ عامل.

درجة النجاح فى العمل	درجة النجاح فى البرنامج التدريبى		
	دون المتوسط	متوسط	أعلى من المتوسط
ضعيف	٢٣	٦٠	٢٩
متوسط	٢٨	٧٩	٦٠
جيد	٩	٤٩	٦٣

حدد الفرض الذى تختبره ثم أجرى الاختبار المناسب عند $\alpha = ٥\%$

٥- لدراسة العلاقة بين طول مدة خدمة العامل فى أحد المتاجر الكبرى وتقبل هؤلاء العاملين لأثر التخفيضات السنوية على حجم المبيعات فقد جمعت البيانات التالية من أفراد عينة عشوائية مكونة من ٤٤٠ عاملاً. حدد الاختبار الإحصائى المناسب ثم استخدمه عند مستوى المعنوية $\alpha = ٥\%$.

التقييم	مدة الخدمة			
	سنتين أو أقل	٣ - ٥	٦ - ١٠	أكثر من ١٠
تزداد المبيعات	٤٨	٥١	٤٣	٣٥
لا تتأثر	٢٢	٤٢	٥٩	٥٧
تنخفض	١٥	٢٢	٣٣	١٣

٦- الجدول التالي يبين توزيع عينة من الأسر بحسب الدخل السنوى وعدد الأطفال بكل:

الدخل السنوى للأسرة بالجنيه	عدد الأطفال			
	صفر	١	٢	أكثر من ٢
أقل من ٥٠٠	١٥	٢٧	٥٠	٤٣
٥٠٠ - ١٠٠٠	٢٥	٣٧	١٢	٨
أكثر من ١٠٠٠	٨	١٣	٩	١٠

يبين الفرض الإحصائى المناسب ثم استخدمه فى اختبار الدخل السنوى وعلاقته بعدد الأطفال عند $\alpha = ٥\%$.

ثم بفرض أن عينة الأسر ذات الدخل أقل من ٥٠٠ جنيه بلغت ١٣٥ أسرة وعينة الأسر ذات الدخل من ٥٠٠ إلى ١٠٠٠ جنيه كانت ١٨٢ أسرة أما عينة الأسر ذات الدخل أكبر من ١٠٠٠ جنيه فكانت مكونة من ٤٠ أسرة ماذا يكون الاختبار ؟

وما هو قرارك عند $\alpha = ٥\%$ ؟

٧- الجدول التالي يبين عدد الحوادث التى وقعت للعمال واللى سجلت فى فترات العمل الثلاث فى أحد الوحدات الإنتاجية:

فترة العمل			
الأولى	الثانية	الثالثة	مجموع
٢٥	٣٠	٤١	٩٦

- هل هناك شك فى عدم وجود اختلاف معنوى بين فترات العمل الثلاث من حيث عدد الحوادث التى تقع عند مستوى المعنوية $\alpha = 0.05$.
- ٨- إذا كان متوسط الوقت الذى تم فيه تجميع الوحدة من منتج ما فى أحد مراكز التجميع هو ٣٢ دقيقة بإنحراف معيارى قدره ٠,١٢٢ دقيقة. ونقضى خطة ضبط الإنتاج فى مركز التجميع أن يكون تباين وقت التجميع داخل الفترة (٠,٠٠٨ ، ٠,٠٣) دقيقة وإلا أوقفت عملية التجميع لإعادة المعايرة. وقد تبين من تسجيل وقت التجميع لعينة عشوائية عددها ٢٥ وحدة أن $\bar{x} = 0.170$ دقيقة. عين التقدير بفترة ثقة $95\% \pm \sigma$.
- ٩- اختبر الفرض بأن $\sigma^2 = 0.015$ دقيقة مقابل الفرض البديل $\sigma^2 < 0.015$ اعتماداً على نتائج التجربة السابق بمستوى معنوية ٠,٠١ .
- ١٠- الجدول التالى يبين نسبة الوحدات المعيبة من إنتاج آلة ما فى ٢٥ يوماً متتالية:

١٠,٠	١٢,٣	٩,٤	١٠,٣	٩,٤	٨,٩	١٠,٣	٨,٢
١٠,٤	٨,٨	١٠,١	١١,٨	١٢,٠	١١,١	١١,٤	٨,٤
١١,٩	١٠,٩	٩,٧	٨,٦	٧,٤	١٢,٢	٩,٩	٩,٣
١١,٢١							

- اختبر عند مستوى المعنوية 0.05% أن نسبة المعيب $<$ الوسيط = نسبة المعيب $>$ الوسيط.
- ١١- البيانات التالية هى طول مدة الانتظار فى محطة الأتوبيس بالدقيقة لأفراد عينة عشوائية مكونة من ١٢ راكباً:
- ٤, ١٠, ٧, ٨, ٧, ٥, ٩, ٧, ٣, ٩, ٧, ٢
- استخدم اختبار الإشارة عند $\alpha = 0.05\%$ لإختبار الفرض $\mu = 5$ دقائق مقابل الفرض البديل $\mu \neq 5$ دقائق ثم أعد الاختبار باستخدام اختبار بارامترى مناسب وعلق على نتائجك.

١٢- فيما يلي المدة (بالدقيقة) التي استغرقها أفراد عينتين عشوائيتين من الرجال والنساء في أداء اختبار للمقابلة للإلتحاق بوظائف أحد المصارف الكبرى:

رجال	٦,٥	١٠,٠	٧,٠٠	٩,٨	٨,٥	٨,٢	٩,٠	٨,٢
	١٠,٩	٨,٧	٦,٧	٨,١	٧,٩	٦,٤	٨,٩	
نساء	٨,٦	٧,٨	٨,٣	٦,٦	١٠,٥	٦,٣	٩,٣	٨,٤
	٩,٧	٨,٨	٩,٩	٧,٦				

يستخدم اختبار مجموع إشارات الرتب U لاختبار أن العينتين من مجتمعين متماثلين عند $\alpha = 0.05$.

١٣- سجل عدد حوادث المرور الأسبوعية التي وقعت عند تقاطع ما في إحدى المدن خلال السنة الماضية فكانت على النحو التالي:

عدد الحوادث	صفر	١	٢	٣	٤	٥	٦
التكرار	١٠	١٢	١٢	٩	٥	٣	١

(أ) ما هو التوزيع الاحتمالي الذي يصف هذه النتائج؟ ولماذا كان اختيارك لهذا التوزيع؟

(ب) اختبر الفرض بأن النتائج المسجلة عن حوادث المرور عن تلك السنة. تتبع في توزيعها التوزيع الاحتمالي الذي اخترته في (أ) من هذا السؤال وذلك عند مستوى المعنوية $\alpha = 0.05$.

١٤- فيما يلي عدد الوحدات المعيبة التي أنتجتها آلة ما في ٢٤ يوم عمل متتالية:

١٥	١١	١٧	١٤	١٦	١٢	١٩	١٧
٢١	١٥	١٧	١٩	٢١	١٤	٢٢	١٦
١٩	١٢	١٦	١٤	١٨	١٧	٢٤	١٣

اختبر عشوائية العينة عند $\alpha = 0.05$.

١٥ - الآتي بيان نتيجة فحص عينات زجاجية متابعة للتلف نتيجة الشحن والنقل

(سليم / تالف : س / ت).

س / س / س / س / ت / س / ت / ت / ت / س

س / ت / ت / س / س / س / س / س / س / س / س / س

ت / ت / ت / ت / س /

اختبر عشوائية نتيجة فحص هذه الرسائل عند $\alpha = 1\%$.

١٦- الآتى بيان قيمة المنفق بالمليون جنيه على بحوث التطوير والتنمية فى ١٥

مؤسسة عامي ١٩٦٢ ، ١٩٧٢ .

اختبر الفرض بأنه لم يحدث تغير في الإنفاق على بحوث التطوير والتنمية

ما بين ١٩٦٢ ، ١٩٧٢ مقابل الفرض البديل أن الإنفاق سنة ١٩٧٢ قد

زاد عما كان عليه سنة ١٩٦٢ عند مستوى المعنوية ٥٪.

٨	٧	٦	٥	٤	٣	٢	١	المؤسسة
١١	٧	٢٠	١٧	٩	١٤	٨	١٢	المنفق سنة ١٩٦٢
١١	١١	٢٢	٢٠	٨	١٧	٩	١٤	١٩٦٢
	١٥	١٤	١٣	١٢	١١	١٠	٩	تابع المؤسسة
	١٧	٢٢	١٣	٥	٨	١١	١٢	المنفق سنة ١٩٦٢
	١٦	٢٧	١٧	١٠	١٠	١٣	١٥	١٩٧٢

١٧- فيما يلي سعر البيع بالقطاعى بالجنيه للوحدة من نوعين مختلفين من سلعة

ما في عينة من منافذ البيع في بلد ما خلال فترة معينة:

٨٤	٨٨	٧٦	٨١	٩٢	٩٠	٨٩	النوع أ
٩٠	٧١	٨٩	٨٧	٨١	٩٣	٧٨	النوع ب
					٨٢	٩٦	

اختبر الفرض بأن النوعين لا يختلفان فى سعر البيع للوحدة بالقطاعى عد
 $\alpha = 5\%$ باستخدام اختبارين لابارامترى وآخر بارامترى.

١٨- فىما يلى المدة (بالدقيقة) التى استغرقت فى إنجاز عملية ما فى أربعة
 فروع مختلفة لأحد المصارف لعينة من العملاء عددهم ٦ عملاء فى كل
 حالة:

الفرع أ	٢٠	٢٥	٣٨	٣١	٢٣	٢٢
الفرع ب	١٨	١٩	١٢	٢١	٢٤	٩
الفرع ج	١١	٨	١٠	٢٧	٢٥	١٤
الفرع د	١٤	١٧	٢٥	٢٨	٢١	١٥

حلل نتائج هذه الدراسة مستخدماً تحليل التباين فى اتجاه واحد باستخدام
 الأسلوبين البارامترى واللابارامترى عند $\alpha = 5\%$ وناقش نتائجك.

١٩- فىما يلى عدد الكيلومترات / جالون التى استهلكت فى اختبار لمتوسط
 استهلاك الوقود لثلاثة أنواع من الوقود (أ، ب، ج).

الفرع أ	٢٨	٢٣	٢٦	٣١	١٤	٢٩
الفرع ب	٢١	٣١	٣٢	١٩	٢٧	١٦
الفرع ج	٢٤	١٧	٢١	٣١	٢٢	١٨

استخدم تحليل التباين فى اتجاه واحد (اختبار هـ) لتحليل نتائج هذه التجربة
 عند $\alpha = 5\%$.

ثم أعد التحليل باستخدام تحليل التباين فى اتجاه واحد (الأسلوب
 البارامترى) عند $\alpha = 5\%$.

ثم ناقش النتائج التى تصل إليها فى الحالتين.

٢٠- لاختبار العلاقة بين الإنفاق على الأمن الصناعى (المنفق بالجنيه / عامل)
 ومعدل حوادث إصابات العمل فى ١١ مؤسسة صناعية فقد جمعت البيانات
 التالية:

المؤسسة	٦	٢	٣	٤	٥	٦	٧
الإنفاق بالجنيه	٦٠	٤٥	٣٠	٢٠	٢٨	٤٢	٣٩
الحوادث	٣	٧	٦	٩	٧	٤	٨
تابع المؤسسة	٨	٩	١٠	١١			
الإنفاق بالجنيه	٥٤	٤٨	٥٨	٢١			
الحوادث	٢	٤	٣	٨			

أوجد قيمة معامل ارتباط الرتب لسبيرمان بين الإنفاق على الأمن الصناعي وإصابات العمل. ثم اختبر معنوية هذا المعامل.

٢١- تعرض البيانات التالية عدد القروض الجديدة التي قدمتها فرعين مختلفين من أفرع أحد المصارف وذلك خلال فترة الـ ١٦ يوماً السابقة:

فرع أ	فرع ب	فرع أ	فرع ب
٨	٨	٦	١٣
٩	٦	٧	٦
٢	١٥	١٨	٩
٦	٤	١١	١٤
٧	١٥	١٠	٦
٧	١٣	٣	١٥
١٠	١٥	١٥	٢
١٠	١٢	١٤	٣

اختبر الفرض العدمي $\mu_a = \mu_b$

$\mu_a < \mu_b$

عند مستوى المعنوية ٥٪.

٢٢- إعتادت أحد الشركات أن تقدم هدايا مختلفة إلى بعض عملائها بداية كل عام ميلادي وذلك تحقيقاً لإستمرارية التعامل مع الشركة. وفي عام ما

قررت الشركة أن تقدم هدية معينة أ (نتيجة حائط تحمل اسم الشركة) إلى
١ عملائها وهدية من نوع آخر ب (طاقم أقلام يحمل الشركة) إلى الثلث
٢ الآخر وأما الثلث الأخير فلم تقدم له هدايا ثم قامت بتصنيف عملائها حسب
نوع الهدايا ودرجة الإستجابة للتعامل مع الشركة وكان النحو التالي:

الهدية			معدل الإستجابة
لا هدية	ب	أ	
٥٥	٥٠	٤٥	زاد
٢٠	١٢	١٥	ثابت
٢٥	٣٨	٤	نقص

استخدم هذه البيانات في اختبار أن نوع الهدية أو عدم تقديم هدية لم يؤثر
على درجة استجابة العملاء مع الشركة عند $\alpha = ٥\%$ وما هو نوع
الاختبار ؟ ثم فسر ما تصل إليه نتائج وما هو القرار الإداري الذي قد
تتصح به في حدود نتائج التحليل ؟

٢٣- تقدم أحد الفرق المسرحية عروضها ٣ مرات في اليوم الواحد على مسرح
صغير ، وفيما يلي عدد المتفرجين في مرات العرض الثلاث في يوم ما
اختير عشوائياً.

العرض الأول	١٧٩	١٦٤	١٨١	١٧٦	١٧٣
	١٧٨	١٦٤			
العرض الثاني	١٦٩	١٧٧	١٦٦	١٧١	١٨١
العرض الثالث	١٧٤	١٦٧	١٦٨	١٨٣	١٦٩

اختبر الفرض (باختبار هـ) عند $\alpha = ٥\%$ بعدم وجود اختلاف في المتوسط
الحقيقي لعدد المتفرجين في مرات العرض الثلاث.

٢٤- فيما يلى إجمالى المنفق من ميزانية الأسرة (س) فى سنة ما بالمائة جنيه وما أنفق على المأكّل (ص) فى ذات السنة بالمائة جنيه لأفراد عينة عشوائية مكونة من ١٠ أسر:

س	٦٥	٦٠	٥٥	٥٠	٤٥	٤٠	٣٥	٣٠	٢٥	٢٠
ص	٣٠	٢٤	٢٥	٢١	٢٠	٢٨	١٦	١٧	١٥	١٤

وإذا كانت قيمة معامل الارتباط الخطى البسيط = ٠,٩٨٠. أحسب معامل ارتباط الرتب لسبيران من هذه البيانات. ثم اختبر الرفض بأن معاملى الارتباط مقدراً بالطريقتين = صفر عند مستوى المعنوية ١٪.

المراجع

أولاً: المراجع العربية :

- ١- الدكتور / عبد اللطيف عبد الفتاح أبو العلا : مدخل إلى الطرق الإحصائية - الطبعة الخامسة (١٩٩٣).
- ٢- الدكتور / عبد اللطيف عبد الفتاح أبو العلا : العينات وتصميم التجارب - الطبعة الثانية (١٩٩٦/١٩٩٧) جامعة المنصورة .
- ٣- الدكتور / عبد اللطيف عبد الفتاح أبو العلا : التحليل الإحصائي ، جامعة المنصورة ، (١٩٩٧-١٩٩٨) .
- ٤- الدكتور / عبد اللطيف عبد الفتاح أبو العلا ، الدكتور / سلطان محمد عبد الحميد : أساسيات الإحصاء (١٩٩٩) مكتبة الجلاء الجديدة - المنصورة .
- ٥- الدكتور / عبد اللطيف عبد الفتاح أبو العلا ، الدكتور / محمد توفيق البلقيني ، الدكتور / سلطان محمد عبد الحميد : التحليل الإحصائي واستخداماته في العلوم التجارية والاجتماعية ، الجزء الثاني ، الطبعة الأولى (٢٠٠١/٢٠٠٢) .
- ٦- الإحصاء والاقتصاد القياسي- سلسلة شوم ترجمة دكتورة سعدية حافظ منتصر (١٩٨٣).
- ٧- الإحصاء في الإدارة - كتاب مترجم دكتور عبد المرضي حامد عزام - دار المريخ ، المملكة العربية السعودية (١٩٩٦) .

ثانياً: المراجع الأجنبية :

1. Anderson, R.L. & Bancroft, T.A. : Statistical Theory In Research. McGraw-Hill Book Company, 1960.

2. C. R. Hicks Fundamental Concepts in the Design of Experiments, 3rd. Fort Worth: Saunders College publishing, 1982
3. Canavos, J., C. and Miller, D.,M. (1999) : Modern Business Statistics, 7th . ed.. International thomson Publishing Company, Duxbury Press, U.S.A
4. Cochrain, W.G., and Cox, G.M.: Experimental designs, John Wiley & Sons, 2nd. ed., 1962.
5. Conover, W.J. : Practical Nonparametric Statistics, John Wiley & Sons Inc., 1971.
6. Conver, W.J., and Iman, R. L., Introduction To Modern Business Statistics, John Wiley &Jons Inc., 1983.
7. Dixon, Wilfrid & Massey, Frank Jr. : Introduction To Statistical Analysis, 4th. ed., McGraw-Hil Book Company, 1983.
8. Fraser, D. A.: Statistics: An Introduction, John Wiley & Sons Inc., 1985
9. Freund, John, E & Williams, Frank, J. & Perles, Benjamin M: Elementary Business Statistics : The Modem Approach, 6th. ed., Prentice-Hall, 1993.
10. Hoel, Poul G. & Jessen, Raymond J.: Basic Statistics For Business And Economics, 2nd. ed., John Wiley & Sons Inc., 1977.

11. J. Neter, W. Wasserman, and M. Kutner, Applied Linear Statistical Models, 2nd ed. Homewood, IL: Richard D. Irwin, 1985.
12. Keller, Gerald & Warrak, Brain : Statistics for Management & Economics, 4th ed., International Thomson Publishing Company, 1997.
13. Kohler, Heinz : Statistics For Business & Economics, 3rd ed., Harper Collins College Publishers, 1994.
14. L. Ott. An Introduction to Statistical Methods and Data Analysis, 4th ed. Belmont. CA: Duxbury Press, 1993.
15. Larsen R. J., and Marx, M. L.,: Statistics, Prentice-Hall International Editions, 1990.
16. Lindgren, B. W.: Statistical Theory, The Macmillan Company, 1962.
17. Maclave, James T. & Benson, George P. & Sincich, : Statistics For Business & Economics, 7th ed., Prentice Hall International Editions, 1998.
18. Mann, Prem S. : Statistics For Business & Economics, John Wiley & Sons Inc., 1995.
19. Mason, R. L. and Lind, D. A. : Statistics Techniques in Business and Economics, 7th ed., Richard D. Irwin Inc., 1990.

20. Mood, Alexander M. & Grabill, Franklin A. :
Introduction To The Theory Of Statistics, McGraw-Hill
Book Company, 2nd . ed., 1963.
21. Ostle, O. : Statistics In Research, Iowa State University
Press, 2nd . ed., 4th Printing 1969.
22. R. B. Miller and D. W. Wichern. Intermediate Business
Statistics: Analysis of Variance, Regression and Time
Series. New York: Holt, Rinehart & Winston. 1977.
23. R. D. Moen, T. W. Nolan and L. P. Provost. Improving
Quality Through planned Experimentation. New York:
Me Graw -Hill, Inc , 1991.
24. Sandy, Robert.: Statistics for Business and Economics,
McGraw-Hill Book Company, International Editions,
Statistics Series, 1990.
25. Selby, S. M., editor : Standard Mathematical Tables,
XVI th. ed., The Chemical Rubber Company, 1958. 2S
Sigel, Sidney : Nonparametric Statistics for behavioral
Sciences, McGraw-Hill Book Company, 1956.
26. Snedecor, george & Cochran, William G. : Statistical
Methods, 6th ed., The Iowa State University press, 1967.
27. Wonnacott, Thomas W., & Wonnacott, Ronald J. :
Introductory Statistics For Business And Economics,
2nd . ed., John Wiley & Sons Inc., 1977.

محتويات الكتاب

رقم الصفحة	الموضوع
(٦-٥)	مقدمة
(٤٩-٩)	الفصل الأول : تحليل التباين وتصميم التجارب
١٠	أولاً : تحليل التباين
١١	(١) اختبارات الفروض بشأن تباين مجتمع ما أو عدة مجتمعات ، توزيع χ^2
١٢	(١-١) اختبارات الفروض بشأن تباين المجتمع σ^2
١٤	(٢-١) اختبارات الفروض بشأن تباين مجتمعين $\sigma_1^2 = \sigma_2^2$ ، توزيع F
١٨	(٢) اختبارات الفروض بشأن $(\mu_2 - \mu_1)$ أو عدة متوسطات
١٩	(٣) تحليل التباين
٢٦	(٤) ملاحظات ختامية
٢٨	ثانياً : تصميم التجارب
٢٨	(١) تعاريف
٢٨	(١-١) التجربة
٢٩	(٢-١) المعالجات
٢٩	(٣-١) وحدة التجربة
٢٩	(٤-١) خطأ التجربة
٣٠	(٥-١) العشوائية

٣١	ثالثاً : التصميم كامل العشوائية
٣١	(١) استخدام عدد متساو من وحدات التجربة في كل معالجة
٣١	(١-١) النموذج الرياضى
٣٣	(٢-١) النموذج الحسابى والتحليل
٣٧	(٢) التصميم كامل العشوائية : عدد غير متساو من وحدات التجربة لكل معالجة
٣٩	(٣) العلاقة بين التصميم كامل العشوائية حيث (ل=٢) واختبار الفرض : $\mu_2 = \mu_1$
٤٢	(٤) المقارنات الفردية
٤٣	(٥) تعليق ختامى
٤٤	تمارين
(٩٦-٥٠)	الفصل الثانى : التصنيف متعدد الاتجاهات
٥١	أولاً : تصميم القطاعات الكاملة العشوائية
٥٢	(١) القطاعات الكاملة العشوائية بنون تكرار
٥٣	(١-١) النموذج الرياضى والنموذج الحسابى
٥٩	(٢-١) طبيعة حد البواقى
٦٠	(٣-١) الكفاءة النسبية للنموذج
٦٥	(٤-١) القراءات المفقودة
٦٨	(٥-١) الخطأ المعيارى الفرق بين متوسطين
٦٩	(٢) القطاعات الكاملة العشوائية مع التكرار
٦٩	(١-٢) مقدمة
٦٩	(٢-٢) النموذج الرياضى

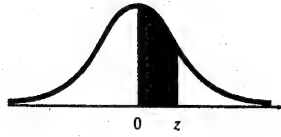
٧١	(٣-٢) نموذج جدول تحليل التباين لتصميم القطاعات الكاملة العشوائية مع التكرار : (النموذج الثابت)
٧٤	ثانياً : المربع اللاتيني
٧٤	(١) مقدمة
٧٧	(٢) النموذج الرياضي والحسابي
٧٨	(٣) الكفاءة النسبية لنموذج المربع اللاتيني
٧٩	(٤) القراءات المفقودة
٨١	ثالثاً : التحليل العاملي
٨١	(١) مقدمة : الأثر الأساسي والتفاعل
٨٣	(٢) التحليل العاملي لتجربة (٢×٢)
٨٤	(١-٢) النموذج الرياضي
٩١	تمارين
(٩٧-١٩٥)	الفصل الثالث : تحليل الانحدار الخطي البسيط
٩٨	(١) مقدمة
١٠٠	(٢) طرق الحصول على خط الانحدار
١٠٩	(٣) خط الانحدار المستقيم
١٢٨	(٤) العلاقة بين ميل خط الانحدار ومعامل الارتباط
١٣٥	(٥) الخطأ المعياري لتقدير معادلة خط الانحدار
١٤٠	(٦) العلاقة بين الخطأ المعياري لمعادلة خط الانحدار ومعامل الارتباط
١٤٤	(٧) معامل التحديد ومعامل الارتباط
١٤٩	(٨) معادلة خط الانحدار من بيانات مبوبة

١٥٥	(٩) الاستدلال الإحصائي عن معالم خط الانحدار
١٦٠	(١٠) تحليل التباين وتحليل الانحدار
١٦٦	(١١) استخدام نموذج الانحدار في عملية التنبؤ بفترة ثقة
١٨٧	تمارين
(١٩٦-٢٥٦)	الفصل الرابع : الانحدار الخطي المتعدد
١٩٧	(١) مقدمة
١٩٨	(٢) فروض نموذج الانحدار
١٩٨	(٣) تقدير معاملات الانحدار الجزئية
٢٠١	(٤) معامل التحديد ومعامل التحديد المعدل
٢٠٤	(٥) الخطأ المعياري للتقدير
٢٠٦	(٦) معاملات الارتباط الجزئية
٢٠٩	(٧) الاستدلال الإحصائي عن معالم خط الانحدار
٢١٣	(٨) اختبار المعنوية الكلية للانحدار : مدخل تحليل التباين
٢١٦	(٩) اختبار مساهمة المتغيرات التفسيرية : مبدأ مجموع المربعات الإضافي
٢٣٢	(١٠) بعض مشاكل استخدام تحليل الانحدار
٢٣٣	١- مشكلة عدم ثبات التباين
٢٣٦	٢- مشكلة الأزواج الخطي
٢٣٧	٣- مشكلة الارتباط الذاتي للبواقي
٢٥١	تمارين

الفصل الخامس : الطرق اللامعلمية	(٢٥٧-٣٣١)
(١) مقدمة	٢٥٨
(٢) توزيع كا ^٢	٢٦١
(٣) اختبار الإشارة ^١	٢٨٧
(٤) اختبار مجموع الرتب (اختبار U) : مان ويتنى	٢٩٦
(٥) اختبارات ولكوكسن التي تعتمد على الرتب	٣٠٤
(٦) تحليل التباين في اتجاه واحد بالرتب / اختبار كروسكال - والس	٣١٢
(٧) معامل ارتباط الرتب لسبيرمان واختبار معنويته	٣١٧
تمارين	٣٢٢
المراجع	(٣٣٥-٣٣٢)
محتويات الكتاب	(٣٤٠-٣٣٦)
الجدول الإحصائية	(٣٥٦-٣٤١)

فهرس الجداول الإحصائية

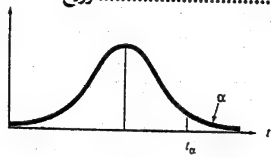
جدول رقم (١)	المساحة تحت المنحنى الطبيعي
جدول رقم (٢)	القيم الأسية
جدول رقم (٣)	توزيع ت
جدول رقم (٤)	توزيع ف ، $\alpha = .10$
جدول رقم (٥)	توزيع ف ، $\alpha = .05$
جدول رقم (٦)	توزيع ف ، $\alpha = .025$
جدول رقم (٧)	توزيع ف ، $\alpha = .01$
جدول رقم (٨)	اختبار ولكوكسن : العينات المستقلة
جدول رقم (٩)	اختبار ولكوكسن : العينات المستقلة
جدول رقم (١٠)	كا ^٢
جدول رقم (١١)	احصاء ديرين — واطسون $\alpha = .05$
جدول رقم (١٢)	احصاء ديرين — واطسون $\alpha = .01$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Source: Abridged from Table I of A. Hald, *Statistical Tables and Formulas* (New York: Wiley), 1952. Reproduced by permission of A. Hald.

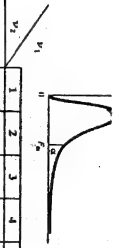
λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
.00	1.000000	2.05	.128735	4.05	.017422	6.05	.002358	8.05	.000319
.05	.951229	2.10	.122456	4.10	.016573	6.10	.002243	8.10	.000304
.10	.904837	2.15	.116484	4.15	.015764	6.15	.002133	8.15	.000289
.15	.860708	2.20	.110803	4.20	.014996	6.20	.002029	8.20	.000275
.20	.818731	2.25	.105399	4.25	.014264	6.25	.001930	8.25	.000261
.25	.778801	2.30	.100259	4.30	.013569	6.30	.001836	8.30	.000249
.30	.740818	2.35	.095369	4.35	.012907	6.35	.001747	8.35	.000236
.35	.704688	2.40	.090718	4.40	.012277	6.40	.001661	8.40	.000225
.40	.670320	2.45	.086294	4.45	.011679	6.45	.001581	8.45	.000214
.45	.637628	2.50	.082085	4.50	.011109	6.50	.001503	8.50	.000204
.50	.606531	2.55	.078082	4.55	.010567	6.55	.001430	8.55	.000194
.55	.576950	2.60	.074274	4.60	.010052	6.60	.001360	8.60	.000184
.60	.548812	2.65	.070651	4.65	.009562	6.65	.001294	8.65	.000175
.65	.522046	2.70	.067206	4.70	.009095	6.70	.001231	8.70	.000167
.70	.496585	2.75	.063928	4.75	.008652	6.75	.001171	8.75	.000158
.75	.472367	2.80	.060810	4.80	.008230	6.80	.001114	8.80	.000151
.80	.449329	2.85	.057844	4.85	.007828	6.85	.001059	8.85	.000143
.85	.427415	2.90	.055023	4.90	.007447	6.90	.001008	8.90	.000136
.90	.406570	2.95	.052340	4.95	.007083	6.95	.000959	8.95	.000130
.95	.386741	3.00	.049787	5.00	.006738	7.00	.000912	9.00	.000123
1.00	.367879	3.05	.047359	5.05	.006409	7.05	.000867	9.05	.000117
1.05	.349938	3.10	.045049	5.10	.006097	7.10	.000825	9.10	.000112
1.10	.332871	3.15	.042852	5.15	.005799	7.15	.000785	9.15	.000106
1.15	.316637	3.20	.040762	5.20	.005517	7.20	.000747	9.20	.000101
1.20	.301194	3.25	.038774	5.25	.005248	7.25	.000710	9.25	.000096
1.25	.286505	3.30	.036883	5.30	.004992	7.30	.000676	9.30	.000091
1.30	.272532	3.35	.035084	5.35	.004748	7.35	.000643	9.35	.000087
1.35	.259240	3.40	.033373	5.40	.004517	7.40	.000611	9.40	.000083
1.40	.246597	3.45	.031746	5.45	.004296	7.45	.000581	9.45	.000079
1.45	.234570	3.50	.030197	5.50	.004087	7.50	.000553	9.50	.000075
1.50	.223130	3.55	.028725	5.55	.003887	7.55	.000526	9.55	.000071
1.55	.212248	3.60	.027324	5.60	.003698	7.60	.000501	9.60	.000068
1.60	.201897	3.65	.025991	5.65	.003518	7.65	.000476	9.65	.000064
1.65	.192050	3.70	.024724	5.70	.003346	7.70	.000453	9.70	.000061
1.70	.182684	3.75	.023518	5.75	.003183	7.75	.000431	9.75	.000058
1.75	.173774	3.80	.022371	5.80	.003028	7.80	.000410	9.80	.000056
1.80	.165299	3.85	.021280	5.85	.002880	7.85	.000390	9.85	.000053
1.85	.157237	3.90	.020242	5.90	.002739	7.90	.000371	9.90	.000050
1.90	.149569	3.95	.019255	5.95	.002606	7.95	.000353	9.95	.000048
1.95	.142274	4.00	.018316	6.00	.002479	8.00	.000336	10.00	.000045
2.00	.135335								



ν	$f_{.100}$	$f_{.050}$	$f_{.025}$	$f_{.010}$	$f_{.005}$	$f_{.001}$	$f_{.0005}$
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.399	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Source: This table is reproduced with the kind permission of the Trustees of Biometrika from E. S. Pearson and H. O. Hartley (eds.), *The Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., Biometrika, 1966.

جدول رقم (4) توزيع ف ، $\alpha = 10$



NUMERATOR DEGREES OF FREEDOM

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.48
3	3.54	3.56	3.57	3.58	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59	3.59
4	2.54	2.56	2.57	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58	2.58
5	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
6	1.78	1.79	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80
7	1.59	1.60	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61	1.61
8	1.46	1.47	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48
9	1.36	1.37	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38
10	1.28	1.29	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30
12	1.18	1.19	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.20
14	1.10	1.11	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12	1.12
16	1.04	1.05	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06
18	1.00	1.01	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
20	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
22	0.95	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
24	0.93	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
26	0.92	0.93	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
28	0.91	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
30	0.90	0.91	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
40	0.84	0.85	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
60	0.75	0.76	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77
120	0.64	0.65	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
∞	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50

Source: From M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution," *Biometrika*, 1943, 33, 73-86.
Reproduced by permission of the Biometrika Trustees.



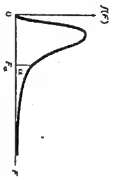
$\alpha = .05$, F_{α} جدول (5)

NUMERATOR DEGREES OF FREEDOM

v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.3	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.47	19.48	19.49	19.50	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	6.71	6.25	6.07	5.92	5.80	5.72	5.66	5.62	5.59	5.56	5.54	5.52	5.49	5.47	5.45	5.42	5.40	5.39	5.38
5	5.19	4.75	4.57	4.42	4.32	4.25	4.19	4.15	4.12	4.09	4.07	4.05	4.02	4.00	3.98	3.94	3.92	3.90	3.89
6	4.35	3.92	3.74	3.59	3.49	3.42	3.36	3.32	3.29	3.27	3.25	3.22	3.19	3.17	3.15	3.12	3.08	3.06	3.05
7	3.78	3.35	3.17	3.02	2.92	2.85	2.79	2.75	2.72	2.69	2.67	2.64	2.61	2.59	2.56	2.53	2.49	2.47	2.45
8	3.36	2.93	2.75	2.60	2.50	2.43	2.37	2.33	2.30	2.28	2.25	2.22	2.19	2.17	2.15	2.12	2.08	2.06	2.04
9	3.02	2.59	2.41	2.26	2.16	2.09	2.03	1.99	1.96	1.94	1.91	1.88	1.85	1.83	1.81	1.78	1.74	1.72	1.70
10	2.75	2.32	2.14	2.00	1.89	1.82	1.76	1.72	1.69	1.67	1.64	1.61	1.58	1.56	1.54	1.51	1.47	1.45	1.43
11	2.54	2.11	1.93	1.79	1.68	1.61	1.55	1.51	1.48	1.46	1.43	1.40	1.37	1.35	1.33	1.30	1.26	1.24	1.22
12	2.37	1.94	1.76	1.62	1.51	1.44	1.38	1.34	1.31	1.29	1.26	1.23	1.20	1.18	1.16	1.13	1.09	1.07	1.05
13	2.24	1.81	1.63	1.49	1.38	1.31	1.25	1.21	1.18	1.16	1.13	1.10	1.07	1.05	1.03	1.00	0.96	0.94	0.92
14	2.13	1.70	1.52	1.38	1.27	1.20	1.14	1.10	1.07	1.05	1.02	0.99	0.96	0.94	0.92	0.89	0.85	0.83	0.81
15	2.04	1.61	1.43	1.29	1.18	1.11	1.05	1.01	0.98	0.96	0.93	0.90	0.87	0.85	0.83	0.80	0.76	0.74	0.72
16	1.96	1.53	1.35	1.21	1.10	1.03	0.97	0.93	0.90	0.88	0.85	0.82	0.79	0.77	0.75	0.72	0.68	0.66	0.64
17	1.89	1.46	1.28	1.14	1.03	0.96	0.90	0.86	0.83	0.81	0.78	0.75	0.72	0.70	0.68	0.65	0.61	0.59	0.57
18	1.83	1.40	1.22	1.08	0.97	0.90	0.84	0.80	0.77	0.75	0.72	0.69	0.66	0.64	0.62	0.59	0.55	0.53	0.51
19	1.78	1.35	1.17	1.03	0.92	0.85	0.79	0.75	0.72	0.70	0.67	0.64	0.61	0.59	0.57	0.54	0.50	0.48	0.46
20	1.73	1.30	1.12	0.98	0.87	0.80	0.74	0.70	0.67	0.65	0.62	0.59	0.56	0.54	0.52	0.49	0.45	0.43	0.41
21	1.69	1.26	1.08	0.94	0.83	0.76	0.70	0.66	0.63	0.61	0.58	0.55	0.52	0.50	0.48	0.45	0.41	0.39	0.37
22	1.65	1.22	1.04	0.90	0.79	0.72	0.66	0.62	0.59	0.57	0.54	0.51	0.48	0.46	0.44	0.41	0.37	0.35	0.33
23	1.62	1.19	1.01	0.87	0.76	0.69	0.63	0.59	0.56	0.54	0.51	0.48	0.45	0.43	0.41	0.38	0.34	0.32	0.30
24	1.59	1.16	0.98	0.84	0.73	0.66	0.60	0.56	0.53	0.51	0.48	0.45	0.42	0.40	0.38	0.35	0.31	0.29	0.27
25	1.56	1.13	0.95	0.81	0.70	0.63	0.57	0.53	0.50	0.48	0.45	0.42	0.39	0.37	0.35	0.32	0.28	0.26	0.24
26	1.54	1.11	0.93	0.79	0.68	0.61	0.55	0.51	0.48	0.46	0.43	0.40	0.37	0.35	0.33	0.30	0.26	0.24	0.22
27	1.52	1.09	0.91	0.77	0.66	0.59	0.53	0.49	0.46	0.44	0.41	0.38	0.35	0.33	0.31	0.28	0.24	0.22	0.20
28	1.50	1.07	0.89	0.75	0.64	0.57	0.51	0.47	0.44	0.42	0.39	0.36	0.33	0.31	0.29	0.26	0.22	0.20	0.18
29	1.48	1.05	0.87	0.73	0.62	0.55	0.49	0.45	0.42	0.40	0.37	0.34	0.31	0.29	0.27	0.24	0.20	0.18	0.16
30	1.46	1.03	0.85	0.71	0.60	0.53	0.47	0.43	0.40	0.38	0.35	0.32	0.29	0.27	0.25	0.22	0.18	0.16	0.14
40	1.38	0.95	0.77	0.63	0.52	0.45	0.39	0.35	0.32	0.30	0.27	0.24	0.21	0.19	0.17	0.14	0.10	0.08	0.06
60	1.30	0.87	0.69	0.55	0.44	0.37	0.31	0.27	0.24	0.22	0.19	0.16	0.13	0.11	0.09	0.06	0.02	0.00	0.00
120	1.22	0.79	0.61	0.47	0.36	0.29	0.23	0.19	0.16	0.14	0.11	0.08	0.05	0.03	0.01	0.00	0.00	0.00	0.00
∞	1.10	0.67	0.49	0.35	0.24	0.17	0.11	0.07	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: From H. M. Thompson, "Tables of Percentage Points of the Inverse Beta (F) Distribution," Biometrika, 1943, 33, 73-88.

27



جدول رقم (١) توزيع ف ، $\alpha = .025$

DENOMINATOR DEGREES OF FREEDOM																			
ν_1	NUMERATOR DEGREES OF FREEDOM																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	12.22	12.44	12.44	12.45	12.46	12.47	12.47	12.48	12.48	12.49	12.49	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50
4	10.13	10.45	10.45	10.45	10.46	10.46	10.46	10.47	10.47	10.47	10.47	10.48	10.48	10.48	10.48	10.48	10.48	10.48	10.48
5	10.01	10.43	10.43	10.43	10.44	10.44	10.44	10.45	10.45	10.45	10.45	10.46	10.46	10.46	10.46	10.46	10.46	10.46	10.46
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.62	3.57	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.34	3.99	3.75	3.59	3.48	3.40	3.33	3.27	3.17	3.07	2.97	2.91	2.86	2.80	2.74	2.68	2.62
14	6.30	4.86	4.24	3.89	3.65	3.49	3.38	3.30	3.23	3.17	3.07	2.97	2.87	2.81	2.76	2.70	2.64	2.58	2.52
15	6.20	4.77	4.15	3.80	3.56	3.40	3.29	3.21	3.14	3.08	2.98	2.88	2.78	2.72	2.67	2.61	2.55	2.49	2.43
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.82	2.72	2.66	2.61	2.55	2.49	2.43	2.37
17	6.04	4.62	4.01	3.66	3.43	3.28	3.16	3.06	2.99	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.26
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.88	2.77	2.67	2.57	2.51	2.45	2.39	2.33	2.27	2.21
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.52	2.46	2.40	2.34	2.28	2.22	2.16
20	5.87	4.46	3.86	3.51	3.28	3.12	3.01	2.91	2.84	2.77	2.68	2.57	2.47	2.41	2.35	2.29	2.23	2.17	2.11
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.54	2.44	2.38	2.32	2.26	2.20	2.14	2.08
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.02
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.72	2.65	2.58	2.48	2.38	2.27	2.21	2.15	2.09	2.02	1.95	1.88
27	5.63	4.24	3.64	3.30	3.08	2.92	2.80	2.70	2.63	2.56	2.46	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.86
28	5.61	4.22	3.62	3.28	3.06	2.90	2.78	2.68	2.61	2.54	2.44	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.66	2.59	2.52	2.42	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

Source: From M. Hartigan and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution," *Biometrika*, 1943, 33, 73-88.

جدول رقم (٨) : اختيار رانكوسن : البيانات المتعلقة

Test statistic is the rank sum associated with the smaller sample (if equal sample sizes, either rank sum can be used).

a. $\alpha = .025$ one-tailed; $\alpha = .05$ two-tailed

$n_1 \backslash n_2$	3		4		5		6		7		8		9		10	
	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U
3	5	16	6	18	6	21	7	23	7	26	8	28	8	31	9	33
4	6	19	11	25	12	28	12	32	13	35	14	38	15	41	16	44
5	6	21	12	28	18	37	19	41	20	45	21	49	22	53	24	56
6	7	23	12	32	19	41	26	52	28	56	29	61	31	65	32	70
7	7	26	13	35	20	45	28	56	37	68	39	73	41	78	43	83
8	8	28	14	38	21	49	29	61	39	73	49	87	51	93	54	98
9	8	31	15	41	22	53	31	65	41	78	51	93	63	108	66	114
10	9	33	16	44	24	56	32	70	43	83	54	98	66	114	79	131

b. $\alpha = .05$ one-tailed; $\alpha = .10$ two-tailed

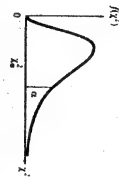
$n_1 \backslash n_2$	3		4		5		6		7		8		9		10	
	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U	T_L	T_U
3	6	15	7	17	7	20	8	22	9	24	9	27	10	29	11	31
4	7	17	12	24	13	27	14	30	15	33	16	36	17	39	18	42
5	7	20	13	27	19	36	20	40	22	43	24	46	25	50	26	54
6	8	22	14	30	20	40	28	50	30	54	32	58	33	63	35	67
7	9	24	15	33	22	43	30	54	34	39	36	41	71	52	84	90
8	9	27	16	36	24	46	32	58	41	71	52	84	54	90	66	95
9	10	29	17	39	25	50	33	63	43	76	54	90	66	105	69	111
10	11	31	18	42	26	54	35	67	46	80	57	95	69	111	83	127

Source: From F. Wilcoxon and R. A. Wilcoxon, "Some Rapid Approximate Statistical Procedures," 1954, 20-23. Courtesy of Lederle Laboratories Division of American Cyanamid Company, Madison, NJ.

جدول رقم (٩) اختبار ولكوكسن : العينات المستقلة

One-Tailed	Two-Tailed	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10
$\alpha = .05$	$\alpha = .10$	1	2	4	6	8	11
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8
$\alpha = .01$	$\alpha = .02$			0	2	3	5
$\alpha = .005$	$\alpha = .01$				0	2	3
		n = 11	n = 12	n = 13	n = 14	n = 15	n = 16
$\alpha = .05$	$\alpha = .10$	14	17	21	26	30	36
$\alpha = .025$	$\alpha = .05$	11	14	17	21	25	30
$\alpha = .01$	$\alpha = .02$	7	10	13	16	20	24
$\alpha = .005$	$\alpha = .01$	5	7	10	13	16	19
		n = 17	n = 18	n = 19	n = 20	n = 21	n = 22
$\alpha = .05$	$\alpha = .10$	41	47	54	60	68	75
$\alpha = .025$	$\alpha = .05$	35	40	46	52	59	66
$\alpha = .01$	$\alpha = .02$	28	33	38	43	49	56
$\alpha = .005$	$\alpha = .01$	23	28	32	37	43	49
		n = 23	n = 24	n = 25	n = 26	n = 27	n = 28
$\alpha = .05$	$\alpha = .10$	83	92	101	110	120	130
$\alpha = .025$	$\alpha = .05$	73	81	90	98	107	117
$\alpha = .01$	$\alpha = .02$	62	69	77	85	93	102
$\alpha = .005$	$\alpha = .01$	55	61	68	76	84	92
		n = 29	n = 30	n = 31	n = 32	n = 33	n = 34
$\alpha = .05$	$\alpha = .10$	141	152	163	175	188	201
$\alpha = .025$	$\alpha = .05$	127	137	148	159	171	183
$\alpha = .01$	$\alpha = .02$	111	120	130	141	151	162
$\alpha = .005$	$\alpha = .01$	100	109	118	128	138	149
		n = 35	n = 36	n = 37	n = 38	n = 39	
$\alpha = .05$	$\alpha = .10$	214	228	242	256	271	
$\alpha = .025$	$\alpha = .05$	195	208	222	235	250	
$\alpha = .01$	$\alpha = .02$	174	186	198	211	224	
$\alpha = .005$	$\alpha = .01$	160	171	183	195	208	
		n = 40	n = 41	n = 42	n = 43	n = 44	n = 45
$\alpha = .05$	$\alpha = .10$	287	303	319	336	353	371
$\alpha = .025$	$\alpha = .05$	264	279	295	311	327	344
$\alpha = .01$	$\alpha = .02$	238	252	267	281	297	313
$\alpha = .005$	$\alpha = .01$	221	234	248	262	277	292
		n = 46	n = 47	n = 48	n = 49	n = 50	
$\alpha = .05$	$\alpha = .10$	389	408	427	446	466	
$\alpha = .025$	$\alpha = .05$	361	379	397	415	434	
$\alpha = .01$	$\alpha = .02$	329	345	362	380	398	
$\alpha = .005$	$\alpha = .01$	307	323	339	356	373	

Source: From F. Wilcoxon and R. A. Wilcox, "Some Rapid Approximate Statistical Procedures," 1964, p. 28.
Courtesy of Lederle Laboratories Division of American Cyanamid Company, Madison, NJ.



جدول رقم (١)

Degrees of Freedom	$\chi^2_{0.99}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.85}$	$\chi^2_{0.80}$	$\chi^2_{0.75}$	$\chi^2_{0.70}$	$\chi^2_{0.65}$	$\chi^2_{0.60}$	$\chi^2_{0.55}$	$\chi^2_{0.50}$	$\chi^2_{0.45}$	$\chi^2_{0.40}$	$\chi^2_{0.35}$	$\chi^2_{0.30}$	$\chi^2_{0.25}$	$\chi^2_{0.20}$	$\chi^2_{0.15}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$	$\chi^2_{0.005}$
1	0.000039	0.000157	0.000983	0.00191	0.002798	0.00364	0.00443	0.00518	0.00589	0.00656	0.00719	0.00778	0.00833	0.00884	0.00931	0.00974	0.01014	0.01051	0.01085	0.01116	0.01144	0.01169	0.01191
2	0.0100023	0.0201007	0.049732	0.073777	0.095906	0.115785	0.133083	0.147790	0.160013	0.170063	0.178329	0.185073	0.190529	0.194926	0.198397	0.201059	0.202996	0.204231	0.204871	0.205021	0.205181	0.205341	0.205501
3	0.0717212	0.114832	0.215795	0.294871	0.354785	0.401541	0.437856	0.465119	0.484419	0.496926	0.503883	0.506456	0.507926	0.508651	0.508991	0.509191	0.509311	0.509391	0.509441	0.509471	0.509491	0.509511	0.509531
4	0.204990	0.297110	0.354785	0.401541	0.437856	0.465119	0.484419	0.496926	0.496926	0.503883	0.506456	0.507926	0.508651	0.508991	0.509191	0.509311	0.509391	0.509441	0.509471	0.509491	0.509511	0.509531	0.509551
5	0.483821	0.675777	0.728795	0.770732	0.803476	0.829443	0.849419	0.864419	0.874419	0.880419	0.883419	0.885419	0.886419	0.887419	0.887926	0.888271	0.888521	0.888691	0.888791	0.888841	0.888871	0.888891	0.888911
6	0.636812	0.871856	0.934785	0.967732	0.990732	1.005732	1.015732	1.021732	1.025732	1.028732	1.030732	1.031732	1.032419	1.032771	1.032926	1.032991	1.033021	1.033041	1.033051	1.033061	1.033071	1.033081	1.033091
7	0.844541	1.134419	1.203419	1.246419	1.274419	1.292419	1.302419	1.308419	1.311419	1.313419	1.314419	1.315419	1.316419	1.316926	1.317271	1.317426	1.317526	1.317591	1.317621	1.317641	1.317651	1.317661	1.317671
8	1.024419	1.344419	1.423419	1.476419	1.504419	1.519419	1.529419	1.535419	1.538419	1.540419	1.541419	1.542419	1.542926	1.543271	1.543426	1.543491	1.543521	1.543541	1.543551	1.543561	1.543571	1.543581	1.543591
9	1.179492	1.534419	1.623419	1.686419	1.724419	1.740419	1.750419	1.756419	1.759419	1.761419	1.762419	1.763419	1.763926	1.764271	1.764426	1.764491	1.764521	1.764541	1.764551	1.764561	1.764571	1.764581	1.764591
10	1.336812	1.734419	1.833419	1.906419	1.954419	1.982419	1.997419	2.007419	2.013419	2.016419	2.018419	2.019419	2.020419	2.020926	2.021271	2.021426	2.021491	2.021521	2.021541	2.021551	2.021561	2.021571	2.021581
11	1.486812	1.934419	2.043419	2.126419	2.184419	2.222419	2.237419	2.247419	2.253419	2.256419	2.258419	2.259419	2.260419	2.260926	2.261271	2.261426	2.261491	2.261521	2.261541	2.261551	2.261561	2.261571	2.261581
12	1.636812	2.134419	2.253419	2.346419	2.414419	2.462419	2.487419	2.497419	2.503419	2.506419	2.508419	2.509419	2.510419	2.510926	2.511271	2.511426	2.511491	2.511521	2.511541	2.511551	2.511561	2.511571	2.511581
13	1.786812	2.334419	2.463419	2.566419	2.644419	2.692419	2.717419	2.727419	2.733419	2.736419	2.738419	2.739419	2.740419	2.740926	2.741271	2.741426	2.741491	2.741521	2.741541	2.741551	2.741561	2.741571	2.741581
14	1.936812	2.534419	2.673419	2.786419	2.874419	2.932419	2.957419	2.967419	2.973419	2.976419	2.978419	2.979419	2.980419	2.980926	2.981271	2.981426	2.981491	2.981521	2.981541	2.981551	2.981561	2.981571	2.981581
15	2.086812	2.734419	2.883419	3.006419	3.104419	3.172419	3.207419	3.217419	3.223419	3.226419	3.228419	3.229419	3.230419	3.230926	3.231271	3.231426	3.231491	3.231521	3.231541	3.231551	3.231561	3.231571	3.231581
16	2.236812	2.934419	3.093419	3.226419	3.334419	3.412419	3.457419	3.467419	3.473419	3.476419	3.478419	3.479419	3.480419	3.480926	3.481271	3.481426	3.481491	3.481521	3.481541	3.481551	3.481561	3.481571	3.481581
17	2.386812	3.134419	3.303419	3.446419	3.564419	3.652419	3.707419	3.717419	3.723419	3.726419	3.728419	3.729419	3.730419	3.730926	3.731271	3.731426	3.731491	3.731521	3.731541	3.731551	3.731561	3.731571	3.731581
18	2.536812	3.334419	3.513419	3.666419	3.794419	3.892419	3.957419	3.967419	3.973419	3.976419	3.978419	3.979419	3.980419	3.980926	3.981271	3.981426	3.981491	3.981521	3.981541	3.981551	3.981561	3.981571	3.981581
19	2.686812	3.534419	3.723419	3.886419	4.024419	4.132419	4.207419	4.217419	4.223419	4.226419	4.228419	4.229419	4.230419	4.230926	4.231271	4.231426	4.231491	4.231521	4.231541	4.231551	4.231561	4.231571	4.231581
20	2.836812	3.734419	3.933419	4.106419	4.254419	4.372419	4.457419	4.467419	4.473419	4.476419	4.478419	4.479419	4.480419	4.480926	4.481271	4.481426	4.481491	4.481521	4.481541	4.481551	4.481561	4.481571	4.481581
21	2.986812	3.934419	4.143419	4.326419	4.484419	4.612419	4.707419	4.717419	4.723419	4.726419	4.728419	4.729419	4.730419	4.730926	4.731271	4.731426	4.731491	4.731521	4.731541	4.731551	4.731561	4.731571	4.731581
22	3.136812	4.134419	4.353419	4.546419	4.714419	4.852419	4.957419	4.967419	4.973419	4.976419	4.978419	4.979419	4.980419	4.980926	4.981271	4.981426	4.981491	4.981521	4.981541	4.981551	4.981561	4.981571	4.981581
23	3.286812	4.334419	4.563419	4.766419	4.944419	5.092419	5.197419	5.207419	5.213419	5.216419	5.218419	5.219419	5.220419	5.220926	5.221271	5.221426	5.221491	5.221521	5.221541	5.221551	5.221561	5.221571	5.221581
24	3.436812	4.534419	4.773419	4.986419	5.174419	5.332419	5.447419	5.457419	5.463419	5.466419	5.468419	5.469419	5.470419	5.470926	5.471271	5.471426	5.471491	5.471521	5.471541	5.471551	5.471561	5.471571	5.471581
25	3.586812	4.734419	4.983419	5.206419	5.404419	5.572419	5.697419	5.707419	5.713419	5.716419	5.718419	5.719419	5.720419	5.720926	5.721271	5.721426	5.721491	5.721521	5.721541	5.721551	5.721561	5.721571	5.721581
26	3.736812	4.934419	5.193419	5.426419	5.634419	5.812419	5.947419	5.957419	5.963419	5.966419	5.968419	5.969419	5.970419	5.970926	5.971271	5.971426	5.971491	5.971521	5.971541	5.971551	5.971561	5.971571	5.971581
27	3.886812	5.134419	5.403419	5.646419	5.864419	6.052419	6.197419	6.207419	6.213419	6.216419	6.218419	6.219419	6.220419	6.220926	6.221271	6.221426	6.221491	6.221521	6.221541	6.221551	6.221561	6.221571	6.221581
28	4.036812	5.334419	5.613419	5.866419	6.094419	6.292419	6.447419	6.457419	6.463419	6.466419	6.468419	6.469419	6.470419	6.470926	6.471271	6.471426	6.471491	6.471521	6.471541	6.471551	6.471561	6.471571	6.471581
29	4.186812	5.534419	5.823419	6.086419	6.324419	6.532419	6.697419	6.707419	6.713419	6.716419	6.718419	6.719419	6.720419	6.720926	6.721271	6.721426	6.721491	6.721521	6.721541	6.721551	6.721561	6.721571	6.721581
30	4.336812	5.734419	6.033419	6.306419	6.554419	6.772419	6.947419	6.957419	6.963419	6.966419	6.968419	6.969419	6.970419	6.970926	6.971271	6.971426	6.971491	6.971521	6.971541	6.971551	6.971561	6.971571	6.971581
40	5.986812	7.334419	7.653419	8.026419	8.374419	8.692419	9.067419	9.077419	9.083419	9.086419	9.088419	9.089419	9.090419	9.090926	9.091271	9.091426	9.091491	9.091521	9.091541	9.091551	9.091561	9.091571	9.091581
50	7.876812	9.534419	9.963419	10.426419	10.874419	11.292419	11.767419	11.777419	11.783419	11.786419	11.788419	11.789419	11.790419	11.790926	11.791271	11.791426	11.791491	11.791521	11.791541	11.791551	11.791561	11.791571	11.791581
60	9.786812	11.634419	12.163419	12.726419	13.214419	13.622419	14.047419	14.057419	14.063419	14.066419	14.068419	14.069419	14.070419	14.070926	14.071271	14.071426	14.071491	14.071521	14.071541	14.071551	14.071561	14.071571	14.071581
70	11.696812	13.734419	14.363419	14.946419	15.454419	15.872419	16.307419	16.317419	16.323419	16.326419	16.328419	16.329419	16.330419	16.330926	16.331271	16.331426	16.331491	16.331521	16.331541	16.331551	16.331561	16.331571	16.331581
80	13.606812	15.834419	16.563419	17.146419	17.654419	18.072419	18.507419	18.517419	18.523419	18.526419	18.528419	18.529419	18.530419	18.530926	18.531271	18.531426	18.531491	18.531521	18.531541	18.531551	18.531561	18.531571	18.531581
90	15.516812	17.934419	18.763419	19.346419	19.854419	20.272419	20.707419	20.717419	20.723419	20.726419	20.728419	20.729419	20.730419	20.730926	20.731271	20.731426	20.731491	20.731521	20.731541	20.731551	20.731561	20.731571	20.731581
100	17.426812	20.034419	20.963419	21.546419	22.054419	22.472419	22.907419	22.917419	22.923419	22.926419	22.928419	22.929419	22.930419	22.930926	22.931271	22.931426	22.931491	22.931521	22.931541	22.931551	22.931561	22.931571	22.931581

Table from C. M. Thompson, "Tables of the Percentage Points of the χ^2 -Distribution," *Biometrika*, Vol. 32, 1945, pp. 199-209, reproduced by permission of the Biometrika Trustees.

n	k = 1		k = 2		k = 3		k = 4		k = 5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	.95	1.54	.82	1.75	.69	1.97	.56	2.21
16	1.10	1.37	.98	1.54	.86	1.73	.74	1.93	.62	2.15
17	1.13	1.38	1.02	1.54	.90	1.71	.78	1.90	.67	2.10
18	1.16	1.39	1.05	1.53	.93	1.69	.82	1.87	.71	2.06
19	1.18	1.40	1.08	1.53	.97	1.68	.86	1.85	.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	.90	1.83	.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	.93	1.81	.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	.96	1.80	.96	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	.99	1.79	.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression. II," *Biometrika*, 1951, 30, 159-178. Reproduced by permission of the *Biometrika* Trustees.

n	k = 1		k = 2		k = 3		k = 4		k = 5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	.81	1.07	.70	1.25	.59	1.46	.49	1.70	.39	1.96
16	.84	1.09	.74	1.25	.63	1.44	.53	1.66	.44	1.90
17	.87	1.10	.77	1.25	.67	1.43	.57	1.3	.48	1.85
18	.90	1.12	.80	1.26	.71	1.42	.61	1.60	.52	1.80
19	.93	1.13	.83	1.26	.74	1.41	.65	1.58	.56	1.77
20	.95	1.15	.86	1.27	.77	1.41	.68	1.57	.60	1.74
21	.97	1.16	.89	1.27	.80	1.41	.72	1.55	.63	1.71
22	1.00	1.17	.91	1.28	.83	1.40	.75	1.54	.66	1.69
23	1.02	1.19	.94	1.29	.86	1.40	.77	1.53	.70	1.67
24	1.04	1.20	.96	1.30	.88	1.41	.80	1.53	.72	1.66
25	1.05	1.21	.98	1.30	.90	1.41	.83	1.52	.75	1.65
26	1.07	1.22	1.00	1.31	.93	1.41	.85	1.52	.78	1.64
27	1.09	1.23	1.02	1.32	.95	1.41	.88	1.51	.81	1.63
28	1.10	1.24	1.04	1.32	.97	1.41	.90	1.51	.83	1.62
29	1.12	1.25	1.05	1.33	.99	1.42	.92	1.51	.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	.94	1.51	.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	.96	1.51	.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	.98	1.51	.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	.95	1.59
35	1.19	1.31	1.14	1.27	1.08	1.44	1.03	1.51	.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression. II," *Biometrika*, 1951, 30, 159-178. Reproduced by permission of the *Biometrika* Trustees.

Non-Parametric Statistics

PROBABILITIES ASSOCIATED WITH VALUES AS LARGE AS OBSERVED VALUES
OF H IN THE KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE BY RANKS

[illegible]

Non-Parametric Statistics

PROBABILITIES ASSOCIATED WITH VALUES AS LARGE AS OBSERVED VALUES OF H IN THE KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE BY RANKS

Sample sizes			H	p	Sample sizes			H	p
n_1	n_2	n_3			n_1	n_2	n_3		
2	1	1	2.7000	.500	4	3	2	6.4444	.008
								6.3000	.011
2	2	1	3.6000	.200				5.4444	.046
								5.4000	.051
2	2	2	4.5714	.067				4.5111	.098
			3.7143	.200				4.4444	.102
3	1	1	3.2000	.300	4	3	3	6.7455	.010
								6.7091	.013
3	2	1	4.2857	.100				5.7909	.046
			3.8571	.133				5.7273	.050
3	2	2	5.3572	.029				4.7091	.092
			4.7143	.048				4.7000	.101
			4.5000	.067	4	4	1	6.6667	.010
			4.4643	.105				6.1667	.022
3	3	1	5.1429	.043				4.9667	.048
			4.5714	.100				4.8667	.054
			4.0000	.129				4.1667	.082
								4.0667	.102
3	3	2	6.2500	.011	4	4	2	7.0364	.006
			5.3611	.032				6.8727	.011
			5.1389	.061				5.4545	.046
			4.5556	.100				5.2364	.052
			4.2500	.121				4.5545	.098
								4.4455	.103
3	3	3	7.2000	.004	4	4	3	7.1439	.010
			6.4889	.011				7.1364	.011
			5.6889	.029				5.5985	.049
			5.6000	.050				5.5758	.051
			5.0667	.086				4.5455	.099
			4.6222	.100				4.4773	.102
4	1	1	3.5714	.200	4	4	4	7.6538	.008
4	2	1	4.8214	.057				7.5395	.011
			4.5000	.076				5.6923	.049
			4.0179	.114				5.6538	.054
4	2	2	6.0000	.014				4.6539	.097
			5.3333	.033				4.5001	.104
			5.1250	.062	5	1	1	3.8571	.143
			4.4583	.100					
			4.1667	.105	5	2	1	5.2500	.036
4	3	1	5.8333	.021				5.0000	.048
			5.2063	.050				4.4500	.071
			5.0000	.057				4.2000	.095
			4.0556	.093				4.0500	.119
			3.8589	.129					